Thesis/
Reports
Kluzek,
E.

~~FINAL REPORT FOR RESEARCH AGREEMENT~~
#INT-94982-RJVA
"Development of a High Resolution Climate Dataset for the Northern Rocky Mountains"
UTAH STATE UNIVERSITY
FS Contact: Dr. William J. Elliot
CoOp Contact: Dr. David Bowles

Development of a High Resolution Climatic Data set for the Northern Rocky Mountains:

Final Report.

Research Joint Venture Agreement INT-94982-RJVA
Agreement #2

Submitted to:
U.S. Department of Agriculture, Forest Service
Rocky Mountain Research Station
Forestry Sciences Laboratory
1221 South Main Street
Moscow, Idaho 83843

Submitted by:
Erik Kluzek, Gail E. Bingham and David S. Bowles

September 20, 1999

UTAH WATER RESEARCH LABORATORY
UTAH STATE UNIVERSITY
Logan, Utah 84322-8200

Thesis/
Reports
Kluzek,
E.

Development of a High Resolution Climatic Data set for the Northern Rocky Mountains:

Final Report.

Research Joint Venture Agreement INT-94982-RJVA
Agreement #2

Submitted to:
U.S. Department of Agriculture, Forest Service
Rocky Mountain Research Station
Forestry Sciences Laboratory
1221 South Main Street
Moscow, Idaho 83843

Submitted by:
Erik Kluzek, Gail E. Bingham and David S. Bowles

September 20, 1999

UTAH WATER RESEARCH LABORATORY
UTAH STATE UNIVERSITY
Logan, Utah 84322-8200

Development of a High Resolution Climate Data set for the Northern Rocky Mountains: Final Report

1. Abstract

This report reports the results of a cooperative research agreement effort to evaluate a procedure for generating high resolution historical and synthetic climate sequences at ungauged locations in the northern mountainous Western United states. A significant problem in managing mountainous areas is the lack of measured climate data that is available to the manager, and the inability to extrapolate the existing data due to terrain variations. The objective of this effort was to determine if meso scale modeling could improve data availability by utilizing the extensive data sets collected daily for weather prediction to supplement or replace the more traditional data sets. The specific objectives were to develop 10 Km simulations of temperature and precipitation over the Northern Idaho region and to evaluate the simulation to determine if this procedure could be used to drive erosion and forest plant models under simulated climate change scenarios.

The climatic simulation data developed in this effort is limited, because most of our effort was spent resolving problems with getting the mesoscale models to operate accurately at this scale. We did extensive analysis on the simulations and found solutions to most of these problems. Once the high resolution modeling problems were resolved, we were able to compare the results with extrapolated (gridded) observed results. Our comparison suggest that scaling accuracy of the input data to the high resolution model limits the accuracy of high resolution model to the point that interpolating observed data provides better results. Our findings show significant limitations in the accuracy of the climate simulations. Although interpolating sparse observed data across high elevation terrain over multiple climatic regions is at best problematic, currently it offers better chance of keeping results in check with actual observations. We concluded that significant improvements will have to be made to nested meteorological regional climate models before surface variables (temperature, precipitation intensity and duration, and heat flux) can be downscaled from globally observed weather data.

2. Introduction

This report summarizes work conducted (Mar/7/1995 to Oct/1/1999) on a research joint venture agreement (INT-94982-RJVA) between the Intermountain Research Station, U. S. Forest Service (USFS), U.S. Department of Agriculture (USDA), and Utah State University (USU). The purpose of the agreement was to develop a high resolution simulated climatic data set for the Northern Rocky Mountain region utilizing the nested meso scale meteorological input model RegCM2 that was developed under our baseline Research Joint Venture Agreement INT-94883-RJVA (Kluzek et. al., 1999). The base objective of this effort was to determine if we could extend the M-CLIGEN effort to study long term climate change effects . M-CLIGEN (Mountain Climate Generator) developed a nested model procedure for generating historical and synthetic climate

sequences at ungauged locations throughout the mountainous Western United States. The driving problem behind this effort arises from the lack of measured climate data available to the Forest Service managers, and the inability to linearly extrapolate the data that does exist due to terrain variations. Our objective in both studies was to mesoscale models to create simulated climate data based on the extensive global weather forecast data sets that have been accumulated since about 1978. The end result of this effort was to try and provide high resolution, elevation mapped temperature and precipitation data sets that could be used to drive the Water Erosion Prediction Project (WEPP) model and to study plant and animal adaptation to projected climate change scenarios. These climate change data sets could then be used by these same models for regional impact assessments by the USFS and other agencies. Unfortunately, as the reports from both projects demonstrate, the simulated surface variables generated using this approach appear to contain larger errors, both in magnitude and bias than data generated by extrapolating the available surface data. Lack of success stems from two general areas, the inadequacy of the regional models available at the time, and the accumulation of error that results from nested simulations using these imperfect models. The original objectives funded under the study were modified to redirect the effort to understand the limitations of the models and to develop simulations for future models.

3. Objectives, History and Resulting Scope

The proposal for this study was written to be an extension of our M-CLIGEN effort, at the time that the regional model, RegCM2 had been debugged and appeared to be working well on the 50 km, 14-year historical regional simulation of the climate of the Western United States. To allow the 50 km data to be scaled to local variables, the M-CLIGEN program included nesting the model for several USFS regions of higher resolution (10 km) and then using a boundary layer model to modify the model surface determined variables to terrain level variables. Data to drive the 10 km model had been organized and the initial model runs made. It became obvious that if this procedure worked, it could provide a significant boost to climate change studies in complex terrain.

In the original proposal the objectives were stated as follows:

1. Develop two 10 km climate scenario grids covering the northern rocky mountain region. One for the 15-year period, 1978 – 1992 and one for one year of doubled CO2 climate. The grid for the existing climate will be validated with SNOTEL data to identify model biases.
2. Compare both the 10 km and the 50 km grids for the existing climate to the double CO2.
3. Compare the 10 km grids to the 50 km grids for both scenarios.

These objectives were very aggressive and were predicated on the success and full funding of the M-CLIGEN effort. In hind sight, the effort proposed exceeded the funding that was available from the Forest Service Climate Change effort, and success required three things to happen together. First, the whole M-CLIGEN model, including the high resolution portion, had to provide an accurate climate simulations without much

additional tuning. Second, the M-CLIGEN program had to continue to be well supported to maintain the infrastructure required to allow the climate change work to piggy back that effort. Third, the German Regional Climate Modeling Simulation program at Max Planck Institut for Meteorologie in Hamburg, Germany had to continue be funded and willing to enter into a cooperative agreement with us to provide the 6 hourly data products from their global climate model ECHAM4.

Unfortunately, problems will both base programs required a redirection of the effort. As we worked with the 10 km simulations under the M-CLIGEN effort, it became clear that there were significant problems with the precipitation distributions being projected by the model. The details of these problems are summarized in our report for project INT-94883-RJVA, Chapter IV, 1999 – which is included in this report as Appendix B. Initially these problems were thought to result from the fact that RegCM2 is a hydrostatic model, which uses a "sigma" surface calculation scheme, one that follows the model terrain grid. It was assumed the problem could be resolved by relaxing the grid to 20 km, so that the between grid vertical rise did not cause result in upslope instability. The problem turned out to be much more significant, and the solution occupied the remaining funding on both the M-CLIGEN and High Resolution efforts. Utah State has invested an additional two years of work in studying and resolving these problems. It should be noted in passing, that the European regional modeling effort encountered similar errors, and as a result, the cooperative effort that we envisioned was never executed.

We met with our Max Planck cooperator, Daniela Jacob from in person on two different occasions in an attempt to establish a cooperative program with them. Our desire was to be able to trade data and experience with them to solve the high resolution problems and begin to look at the doubled CO_2 issues. However, at that time Max Planck was not releasing their data to outside groups until they had finished their own analysis. Max Planck only gives their data to selected collaborators and because we lacked sufficient funding on the M-CLIGEN effort to open a formal simulation cooperative agreement, our request to develop a joint program was turned down. We argued that collaboration with us would provide extensive analysis of their simulations over a different part of the globe. We continued contact with Daniela via e-mail for quite some time. Finally, Daniela finished her 5-year tenure at Max Planck and in 1996 she moved to a new program.

After this proposal was funded, we found that the long-term precipitation being calculated by the 10 km simulation was grossly under-predicted. It was critical that we resolve this problem before simulations over the Northern Rockies would be useful. The FY-1993 M-CLIGEN report points out some of the theoretical problems with mesoscale modeling at 10 km resolution. When we relaxed the grid spacing to 20 km, and found similar results, it became apparent that our difficulties included some of the same problems that had been pointed out in other research. Due to the extensive effort spent on resolving the problems with high-resolution simulations, we could not complete the long term high resolutions simulations that we had originally proposed. Because the solution to the high-resolution simulations problems greatly increased the CPU requirements for these runs, it was determined that completing long-term 10 km simulations was not cost effective. However, two shorter simulations on the region have

been completed, to verify that the solutions developed using data for Northern Utah are applicable to the Northern Idaho area.

4. *High Resolution Mesoscale Model Effort*

To understand the High Resolution effort, it is important to be familiar with the base program, M-CLIGEN, and the long term regional climate simulation performed under that effort. This is critical, because both the tools and the input (50 Km resolution) data for the high resolution simulation are drawn from that effort. To set the stage for the high resolution (20 Km) simulation, we first review the input data set developed under M-CLIGEN.

4.1 Accuracy of the nested simulation in the Northern Rocky Mountain area.

The performance of the 50km and 20km simulations are tied. The 50km simulation is used as input to the 20km simulation. And, as such, the gross domain averaged features of the 20km simulation will be close to the performance of the same area for the 50km simulation. Although the 20km simulation can respond more accurately to the underlying fine scale orography, the domain averaged features can not deviate strongly from the input conditions. In other words, since the 50km simulation is used as input to the 20km simulation the biases of the 50km simulation will appear in the 20km simulation. For example, if the 50km simulation brings in storms at twice the observed frequency the 20km simulation will have no choice but to create storms at this frequency. In the following we show some of the analysis we applied to the 50km simulation for the Northern Rockies area. There are several deficiencies in the 50km simulation that will create similar deficiencies in the high resolution simulations. Summertime precipitation is too high, and precipitation in the Washington-Idaho area is always too high. The model simulates precipitation only slightly better than simple climatology as a forecast. Finally, although the wet-spell length of storms is modeled fairly well for both summer and winter, the dry-spell length between storms is too short in the model. Thus, the model creates storms at too high a frequency.

Some of the problems with summertime convective precipitation may be corrected by the high resolution model that is better able to resolve thunderstorm cells. And the precipitation intensity distribution may be modified by the high resolution simulation as well. However, overall the large scale features of the high resolution simulation will be derived directly from the input conditions. To illustrate some of the inaccuracies with the 50km regional model simulation we examine three areas in the Northern Rockies. In Figure 1, we display the detailed terrain of the western U.S. as well as these analysis regions. Additionally, each region displays the monthly averaged precipitation climatology for that region. The three regions of interest in our case are: Washington-Idaho (WA-ID), Idaho-Montana (ID-MT), and central Idaho (ID). Note that the WA-ID region is in a rain-shadow region from the flow passing over the Cascades and is

significantly drier than the other area's. Also all regions display relatively low precipitation in the summer.

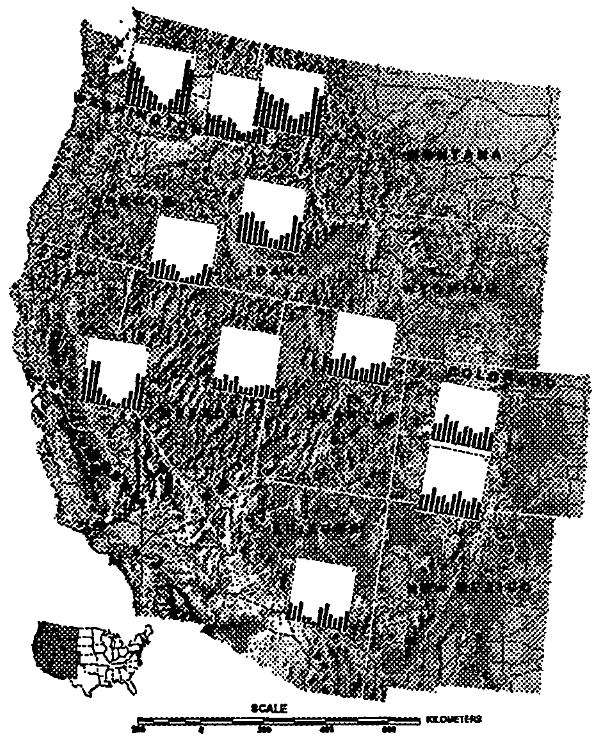
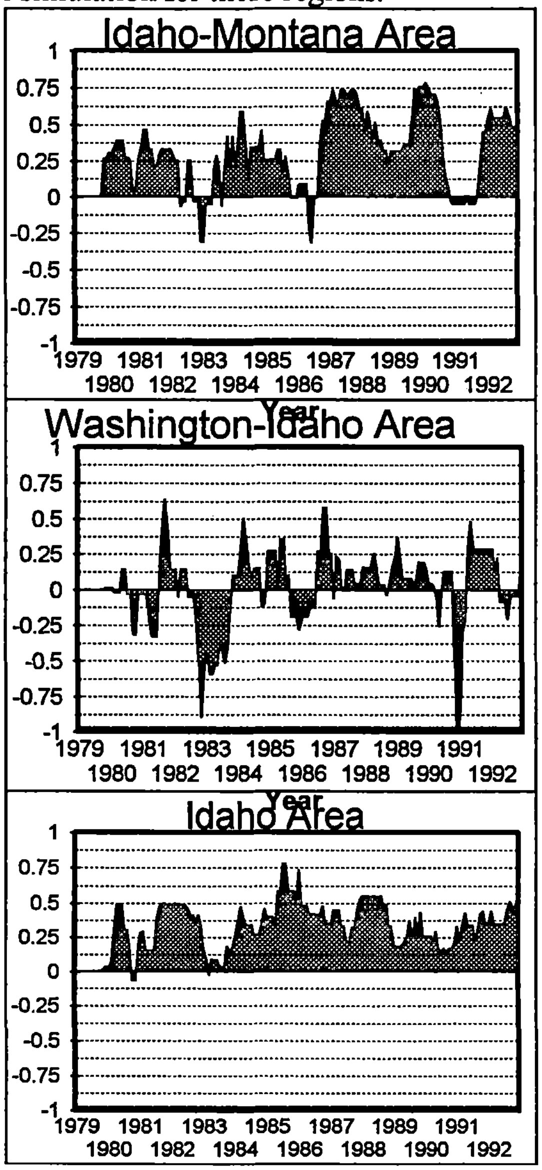


Figure 1 The domain of the M-CLIGEN 50 Km regional simulation showing the three Northern Rocky Mountain areas of detailed analysis studied in that effort.

In the M-CLIGEN effort, Kluzek et. al., 1999, developed a method for gridding the observed data on the same scale as the modeled data and made detailed comparisons of the two data sets. In Figure 2, we see the observed and modeled precipitation climatology for each month for the three analysis areas. Precipitation is matched well for most of the year for the ID, and ID-MT regions. However, summertime precipitation is much too strong for all regions, and the WA-ID region has too much precipitation year-round. Part of the problem with the WA-ID region may be the lack of the 50km terrain for resolving the cascade mountain range, thus the 50km simulation doesn't sense the rain shadow region. The summer precipitation problem becomes more severe as the latitude decreases, becoming the dominant error in the higher sun regions where most of the annual precipitation is derived from summer thunderstorms. The inability to handle summer precipitation is a significant portion of the high resolution simulation problem.

In Figure 3, we display the bias corrected precipitation skill score. This score compares the quality of the model simulation to a forecast using simple observed climatology. Near zero the model is no better than climatology, less than zero it is worse, and near one the model is perfect. The model scores are very low for all three regions, WA-ID is especially low with no apparent improvement over climatology. The other two regions only show moderate skill values. Since the skill is computed accounting for the long-term

model bias this shows the low amount of confidence that can be held with the 50km model simulation for these regions.



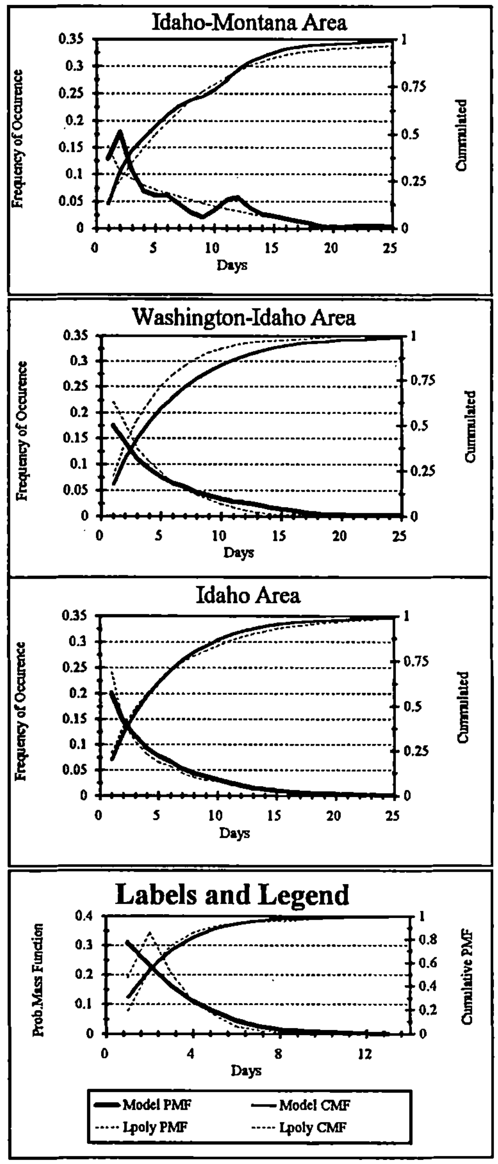


Figure 2. The model precipitation skill score for the three Idaho analysis areas. When the skill score is 1, modeled and observed data agree perfectly. At a skill score of zero, the model is predicting precipitation at the same level as the climatology. Details of the analysis are presented in Kluzek, et al., 1999.

Figure 3. The observed wet spell length distribution compared to the model for the winter period (December, January, and February) for the three Idaho analysis areas. Cummulated frequency is normalized to 1. For more detail see Kluzek, et. al., 1999.

Figure 3 also shows the modeled and observed wet-spell length frequency (length in days that a storm event lasts) for wintertime, while Figure 4 shows the summertime

frequencies. In both cases the model matches the observed fairly well. Although, the WA-ID region is not modeled nearly as well as the other two regions.

Figure 5 shows the summertime dry-spell frequency of model and observed for the three regions. Here, we note that the model is creating precipitation at much too high a frequency, hence the problem with excessive summertime precipitation. Since, much of this problem may be with the convective precipitation parameterization, it's possible, that the high resolution simulation may ameliorate the summertime precipitation problem.

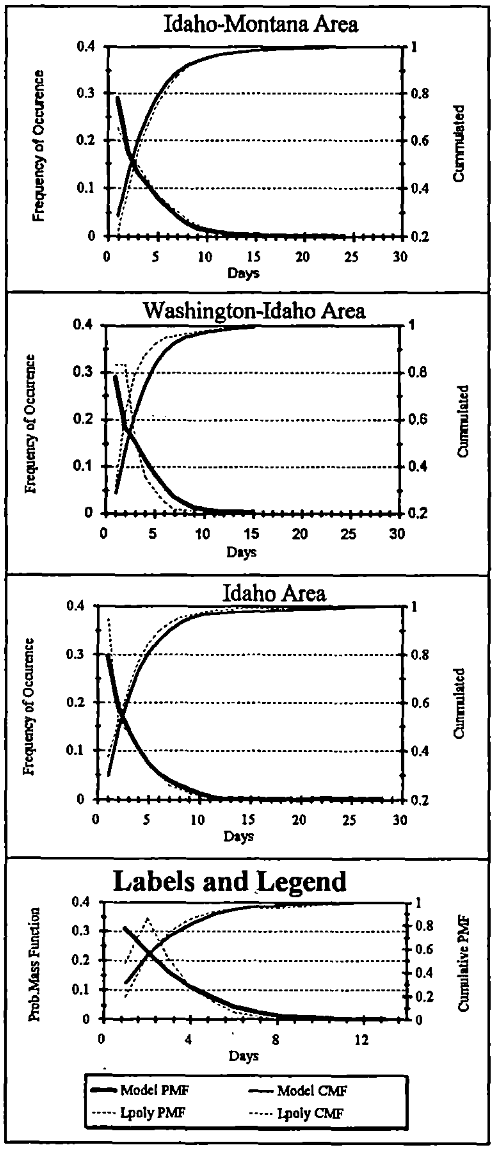


Figure 4. The observed wet spell length distribution compared to the model for summer period (June, July, and August) for the three Idaho analysis areas. Cumulated frequency is normalized to 1. (From Kluzek et. al., 1999)

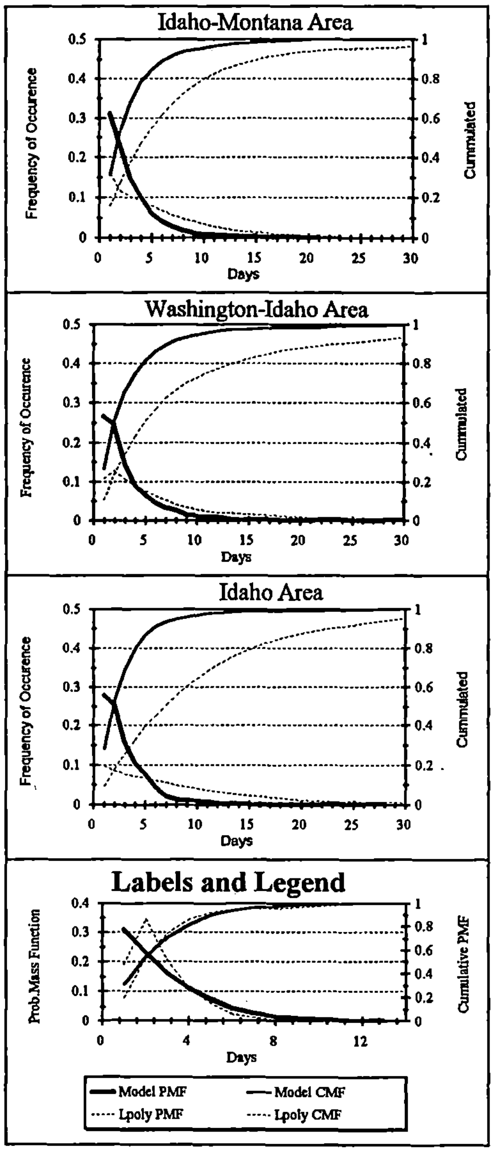


Figure 5. The observed dry spell length distribution compared to the model for summer period (June, July, and August) for the three Idaho analysis areas. Cummulated frequency is normalized to 1. (from Kluzek et. al., 1999)

In summary, a significant and known difference exists in the precipitation data driving the high resolution simulation. A bias also exists in the surface level temperature data, with the monthly mean values in good agreement, but with smaller daily excursions than are predicted by the extrapolated observed data. Note, that this is a comparison of two models, and not a comparison to actual values, but the gridded observed data appears to be more realistic. Storm phasing and location is also different between the two data sets.

4.2 High Resolution Simulations.

Details of the high resolution simulation process can be found in Kluzek et. al., 1999. In summary, the grid scale of the high resolution simulation was expanded to 20 Km to minimize computer time and to eliminate grid scale error resulting from implied vertical motion due to the sigma surfaces in the model. That is, we did not want to be confused by effects resulting from operating the hydrostatic model at a lower resolution than its structure could tolerate in this complex terrain. Our initial examination of results of the high resolution simulations of the Utah area seemed to be providing good quality data. Daily or sub-daily temperatures and precipitation were at reasonable levels. However over longer periods we noticed severe problems with the simulations. Precipitation amounts were too low over the whole domain while over mountain upslopes, they seemed oddly too high. We spent quite a bit of effort to quantify this problem. In the appendix we reproduce a chapter from the M-CLIGEN climate modeling final report (Kluzek et. al., 1999) which goes into extensive detail on this analysis. We will only present the basic information here.

A Limited Area Model (LAM) is a very complex scientific tool. When problems arise, it can be very difficult to determine the cause and, as is often the case, very difficult to find solutions. The LAM simulation is dependent on the following factors; the input lateral boundary conditions (LBC), the lateral boundary relaxation scheme, and finally, on the model and model physics itself. Because the 50 km simulations were successful, the most likely problems were either the input, or a scale dependent problem with the physics. An outline of the areas to examine follows.

- Analyze the input data and compare to observations.
- Examine possible theoretical problems due to scale issues.
- Perform experiments to verify our analysis
- Apply to the Northern Rockies Area

In Figure 6 we see the observed and modeled precipitation for September 1981 over the 10 km model domain. The modeled precipitation is much too low over the entire domain. The distribution of precipitation is reasonably accurate, but there is less precipitation over the lee-side of the Uinta mountain range than observed.

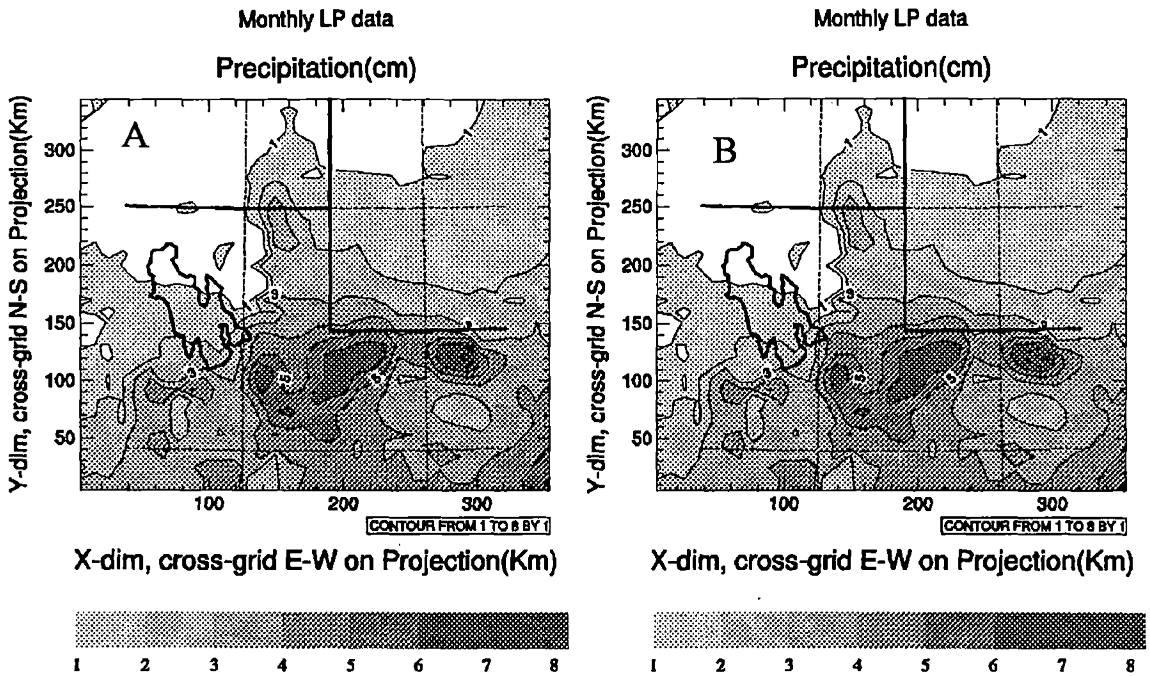


Figure 6. Observed (A) and modeled (B) precipitation for September, 1981 for the Utah area high resolution (20 Km) test case, using original model code and input.

Kluzek et. al., 1999 (see Appendix) goes into in depth analysis of the above, so we will only present out conclusions here. After an in depth analysis and experimentation with RegCM2 we found that successful high resolution simulations require the following:

- Surface pressure was off by 80mbar
- Frequency of boundary updates needs to be increased from every six hours to every hour.
- Cloud water and rain water needs to be input on the lateral boundary conditions.
- Some adjustment to the cloud microphysics may be necessary.

The appendix deals with the last three issues, however, the first problem was not diagnosed at the time that this section was developed, but was included in the Idaho simulation and the following comparison.
Overall, precipitation values become much more reasonable, as shown in Figure 7. However we still note several important deficiencies in the model simulation: precipitation is too low near the lateral boundaries, and there is a point of unrealistically high precipitation (50 cm). The analysis done in the appendix suggests that a possible solution for these problems is a faster update frequency, and the input of large scale cloud and rain water.

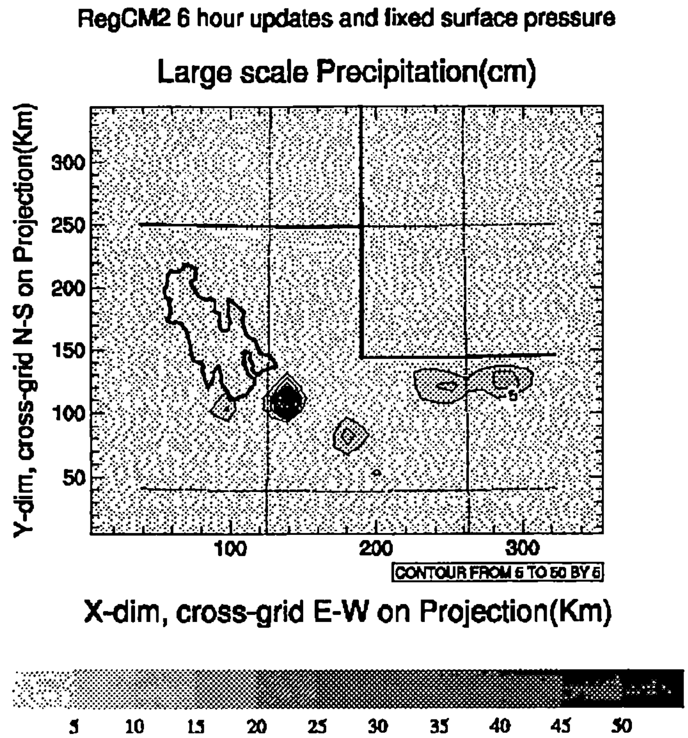


Figure 7. Utah area high resolution precipitation for September, 1981 after correcting for the surface pressure problem. Compare with Figure 6.

4.3 Perform simulations for the Northern Rockies area and compare to observed data.

Figure 8 shows the surface elevation (in meters) of the RegCM2 model domain for the 20 Km Northern Rockies grid. The area of the simulation runs from eastern Washington, through Idaho and into western Montana. The north to south extent runs from just north of the U.S. border to the lower third of the State of Idaho. The northern most region is fairly smooth, with the highest terrain peaking in the southern most part of the grid, in Idaho, at an elevation of 2200 m. The terrain elevation differences are similar to those utilized in the Utah study, but the input air flow is significantly different in the extent two which the input zonal air flow is exposed to up wind orographic influences.

In Figure 9 we see observed and modeled precipitation for September, 1981 for the Northern Rocky Mountain area. The precipitation patterns and amounts from the observed gridding model (Local Polynomial, See Kluzek et. al., 1999) and the high resolution nested simulation are significantly different. Most of the observed precipitation is confined to the high elevation features in northwestern Montana. The model simulation is significantly wetter, and has the same problems observed in the Utah simulations (see Kluzek et. al., 1999).

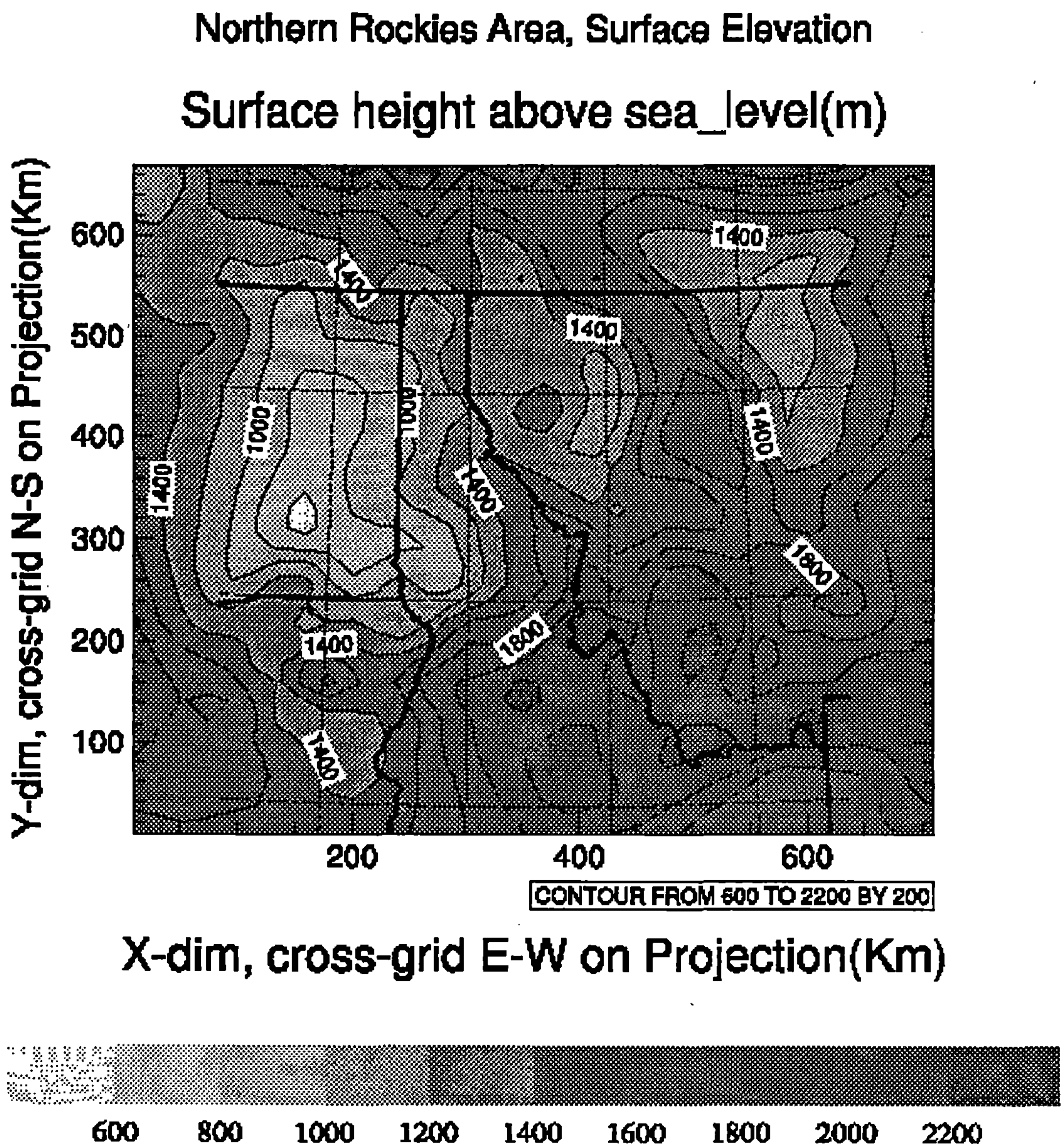


Figure 8. Surface elevation for the Northern Rocky Mountain area used in this climate simulation.

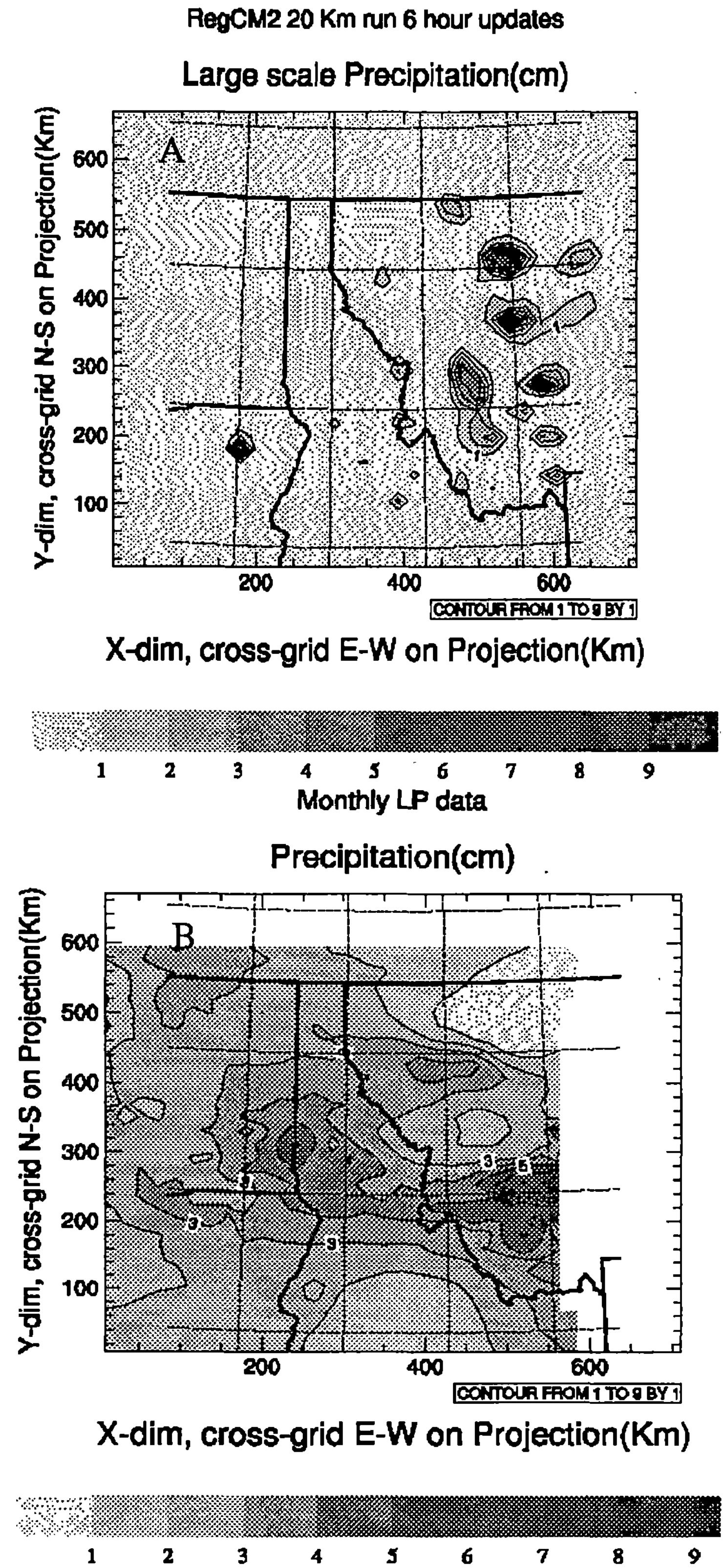


Figure 9. Observed (A) and modeled (B) precipitation (cm) for the Northern Rocky Mountain area for September, 1981.

5. Summary and Conclusions

Through extensive analysis we have been able to find solutions such that our mesoscale model will start to give reasonable answers for high resolution simulations. High resolution simulations require explicit cloud and rain microphysics, one hour updates of lateral boundary data, as well as the inclusion of cloud and rain water on the lateral boundaries. We demonstrated these improvements for both the Northern Utah and Northern Rockies regions. However, although the answers are reasonable, for the Northern Rockies region they are not very useful. The 50km model's wet bias in summertime precipitation is quite large, and the model severely over predicts precipitation for all months over the Washington and Idaho region. Although we believe the high resolution simulations will ameliorate these problems to some extent -- the high resolution simulation can not by nature deviate too wildly from it's input.

As a solution to this problem it would seem natural to apply 4D data assimilation and nudge the model solution to observed states. This may be a necessary technique in order to get simulations of high enough accuracy to use for impact studies. However, we wish to point out caution in this area. 4D assimilation works best when the model is already near the observed state, the further away the models preferred solution is from observed, the less likely that data assimilation will yield good results. By running without any data assimilation we have shown that the models preferred state diverges significantly from observed. Also if the ultimate purpose of study is to simulate unobserved data points, we bring into question the validity of the unobserved data when we can only compare to observed stations that were used to drive the model in the first place.

In order to successfully use mesoscale models it may be necessary to drive models over a very large area at a very fine resolution. It then may be necessary to run with two-way nesting to focus on limited fine mesh areas. Alternatively assimilating data into model output to use as input for high resolution simulations may be necessary for accurate long-term climate simulations. Although these methods are available they would be extremely costly in terms of super-computer time, and development and management costs. The methodology we used was substantially simpler and not as costly, but as we have seen could not provide accurate enough a solution to be used with impact modeling.

In order to simulate historical conditions it seems the most reliable method is the method we used to evaluate the model -- local polynomial regression. This method works reasonably well on daily temperatures and monthly precipitation. Other variables would need to be derived from the foregoing list. By supplementing the local polynomial error estimate with a parameter sensitivity (run the local polynomial changing several basic parameters, then assessing the variability in the solution) we believe reasonable error estimates can be derived. We have seen that the mesoscale model bias is larger than these error estimates and until they can be significantly reduced dependence on the local polynomial should continue. In many cases the error in the local polynomial regression may be too high to be useful, but at least estimates can be performed for values within the error tolerance.

6. References

Kluzek, Erik, Gail E. Bingham and David S. Bowles. 1999. Development of Mountain Climate Generator and Snowpack Model for Erosion Predictions in the Western United States using WEPP: Final Report – Climate Modeling and Data Analysis. Research Joint Venture Agreement INT-94883-RJVA. Submitted to: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Forestry Sciences Laboratory, 1221 South Main Street, Moscow, Idaho 83843, by Utah Water Research Laboratory, Utah State University, Logan, UT. 84322-8200. September 20.

APPENDIX

THE IMPACT OF LATERAL BOUNDARY CONDITIONS ON REGIONALCLIMATE MODELING: A HIGH RESOLUTION TESTCASE IN COMPLEX TERRAIN[1]

Abstract. Recently there is interest in integrating regional climate models at higher resolutions and for longer periods than in the past. In some cases there is considerable work required before a model can produce successful simulations, and in other cases the problems are unique to the resolution. Using the regional model RegCM2, we found the domain average precipitation to be lower than observed and precipitation too tightly bound to mountain up-slopes. Although some of these problems may be unique to RegCM2 or our domain, we believe their solutions are linked to regional modeling at 10 km resolution. We suggest that at this resolution regional models require explicit cloud and rain water microphysics with some modifications. For small domains, cloud and rain water should be input on the lateral boundary conditions and update frequency may need to be increased. We present and discuss extensive analysis of our simulations.

Introduction

There is continued interest in high-resolution climate analysis for various impact studies including hydrological studies of local floods and droughts, and studies of vegetation, forest, and crops. With increased concerns of long-term, large-scale weather patterns such as El-Nino Southern Oscillation (ENSO) and climate change, there is even more concern with the regional impact from large-scale climate variation. Researchers and decision makers want to resolve not only what happens on a global scale, but how large-scale changes affect a particular country, state, political district, or watershed. There are also concerns with atmospheric chemistry and pollutants, the impact of an increasing population, and the feedback a changing climate may have. With new legislation on the global release of greenhouse gases, there will continue to be interest in determining how these changes affect the local climate. There is also a concern that many of the stresses that limit plant growth disappear in global climate model (GCM) studies because of climatic averaging [*Paoli,* 1994]. With computing speeds continuing to increase, GCMs are being run at higher and higher resolutions. Likewise, limited area model (LAM) studies must also continue to be run at higher and higher resolutions. *Kato* [1997] points out that numerical weather prediction (NWP) models are likely to be run operationally as fine as 10 km in the near future. He also points out that prior to this event the modeling community needs to "determine the optimal dynamical framework and precipitating scheme." In the case of LAMs, "dynamical framework" refers to the method of including LBCs [*Waldron et al.,* 1996]: spectral methods,

---

[1] Section from "Development of Mountain Climate Generator and Snowpack Model for Erosion Predictions in the Western United States using WEPP: Final Report – Climate Modeling and Data Analysis, 1999. Research Joint Venture Agreement INT-94883-RJVA, Coauthored by E.B. Kluzek, G.E. Bingham and D.S.Bowles.

continuously varying grid-size, Davies nudging [*Davies and Turner*, 1977], LBC from outer model, plus large-scale forcing internally from the outer model, or two-way nesting. Precipitating scheme refers to the method of solving for precipitation. There are many choices for this (implicit vs. explicit cloud and rain water, as well as sub-grid-scale convective parameterization). The same need to determine the "optimal dynamical framework and precipitating scheme" also applies to regional climate models.

The mechanics of nesting LAM beneath higher resolution GCM runs need to be worked out. A particular challenge for determining the precipitation scheme is that the solution depends on resolution [*Molinari and Dudek*, 1992]. *Molinari and Dudek* [1992] point out that there is not an obvious choice for the precipitation scheme at 10 km resolution. This underlines the need for more experimentation at this level. Regional climate models are good tools for bringing fine-scale information from a GCM simulation. The modeling community is now at the point where there is an interest to run a LAM on the interannual timescale at finer resolution than has been tried before. This study examines some of the scientific problems encountered when a LAM, traditionally run at resolutions on the order of 50 km, is used at a horizontal resolution of 10 km. The most significant error, and perhaps the most important, is the overprediction of precipitation on mountain up-slopes, coupled with an underprediction of precipitation elsewhere [*Nickerson et al.*, 1986; *Zhang et al.,* 1988]. We present analysis of the problem for our simulations and propose possible sources for a solution. We propose that certain deficiencies exist in the RegCM2 explicit cloud microphysics. We also propose that cloud and rain water need to be input on the LBCs for small domains. Finally, we present some problems that occur when LBCs are not updated at a high enough frequency. More precisely, we use the National Center for Atmospheric Research's (NCAR) LAM climate model Regional Climate Model version 2 (RegCM2) [*Giorgi et al.*, 1993a, 1993b]. We do this at 10 km resolution over the complex terrain of the northern Utah Rocky Mountains. See Figure 1 for a plot of our 50 km model grid, and Figure 2 for the nested 10 km model grid. The 10 km grid includes the Great Salt Lake, the Uinta, Wasatch, and Bear River mountain ranges, and parts of Idaho and Wyoming. Also included is the Salt Lake City rawinsonde that is near the center of the domain. The results shown are for one month, September 1981. Further simulations will need to be performed to verify the conclusions of this study for other areas and seasons. We provide the theoretical underpinning for our results based on basic physics, and we believe the results can have validity outside the area and season demonstrated.

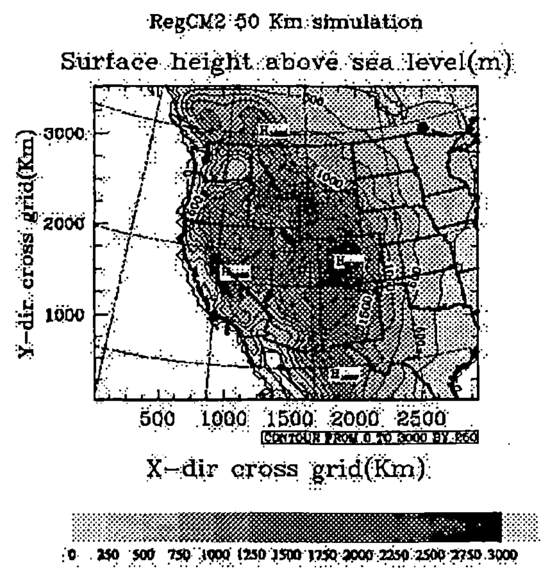


Figure 1. RegCM2 50 km terrain contour map of the western US. Compare these smoothed profiles with the DEM generated terrain.

Discussion

The RegCM2 [*Giorgi et al.*, 1993a, 1993b] model is a hydrostatic, compressible, sigma layer, primitive equation atmospheric model. The basic model is a mesoscale model version 4 (MM4) as described by *Anthes et al.* [1987]. It has sophisticated packages for surface physics (Biosphere Atmosphere Transfer Scheme version 1e (BATS1E)), [*Dickinson et al.*, 1993], lakes [*Hostetler and Giorgi*, 1992], and a non-local atmospheric boundary layer [*Holtslag et al.*,

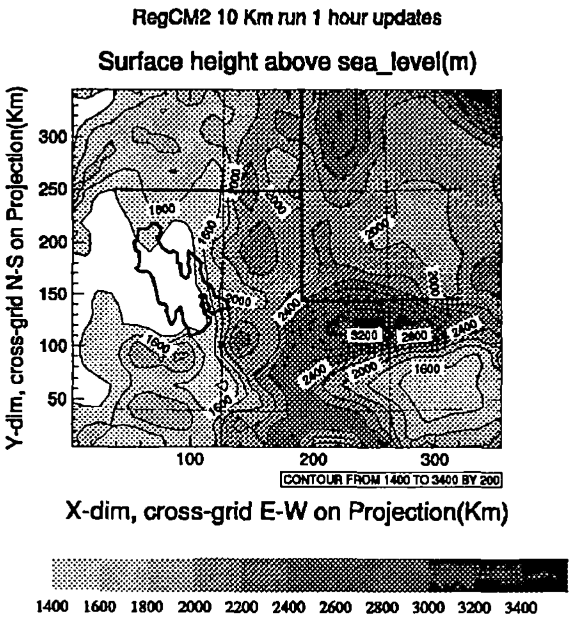


Figure 2 . Surface heights for the 10-km model domain.

1990; *Holtslag and Boville*, 1992]. A delta-Eddington approach to solar radiation [*Briegleb*, 1992] and an option to use the Grell convective parameterization scheme [*Grell et al.*, 1994] – which we used with a *Fritsch and Chappell* [1980] type closure – were included. RegCM2 is an improved version of the RegCM model used in several earlier studies [*Giorgi and Bates*, 1989; *Giorgi*, 1990, 1991; *Medina*, 1991; *Giorgi et al.*, 1992; *Matthews et al.*, 1992].

To save CPU time, RegCM2 replaced the Brown-Campana time differencing approach used in MM4 [*Anthes et al.*, 1987] with a split-explicit approach [*Madala*, 1981]. In this study, two model domains are used – a 50 km model domain and a nested 10 km domain. The first model domain (Figure 1) is similar to the area over the western US studied by *Giorgi et al.* [1992], except that a grid domain of 72 x 60, 50 km square boxes, with 17 vertical layers is used. We apply terrain smoothing on the lateral boundaries, similar to *Hong and Juang* [1998]. The lowest model layer is approximately 20 m off the surface and the top model layer is 80 hPa. In the second model domain (Figure 2) the grid size is 36 x 37, 10 km grid-squares with 20 vertical layers. Like the 50 km simulation, we apply terrain smoothing on the lateral boundaries. For these simulations we use the explicit cloud and rain water scheme contained in RegCM2 [*Hsie and Anthes*, 1984].

In Figure 3 and Figure 4, we show the gridded observed precipitation and error

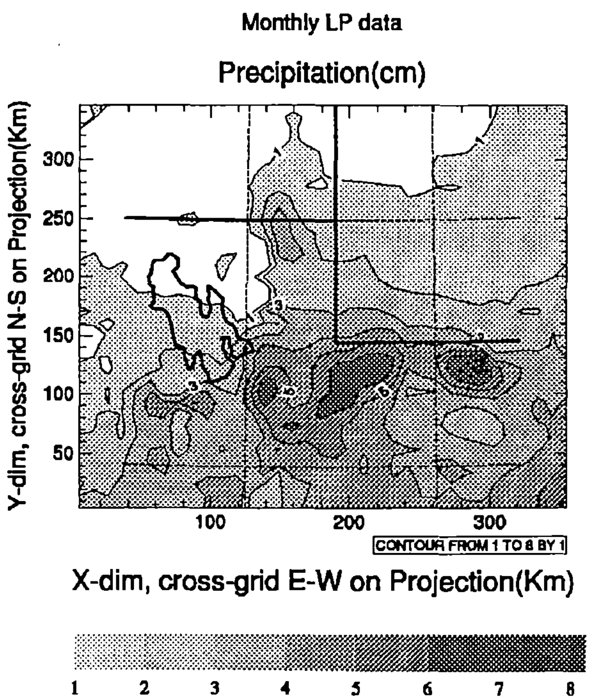


Figure 3. Observed precipitation over model domain for September 1981. Monthly station data

were gridded to the model grid using the local polynomial regression technique.

over the 10 km grid domain for September 1981. The observed precipitation is gridded using a local polynomial regression in latitude, longitude, and elevation [*Lall and Bosworth*, 1994]. This is the same methodology as used in the study for the 50 km data. In this case monthly station data are used to generate the data set, rather than gridding daily data and accumulating the monthly total. Since monthly precipitation is much more smoothly distributed over the

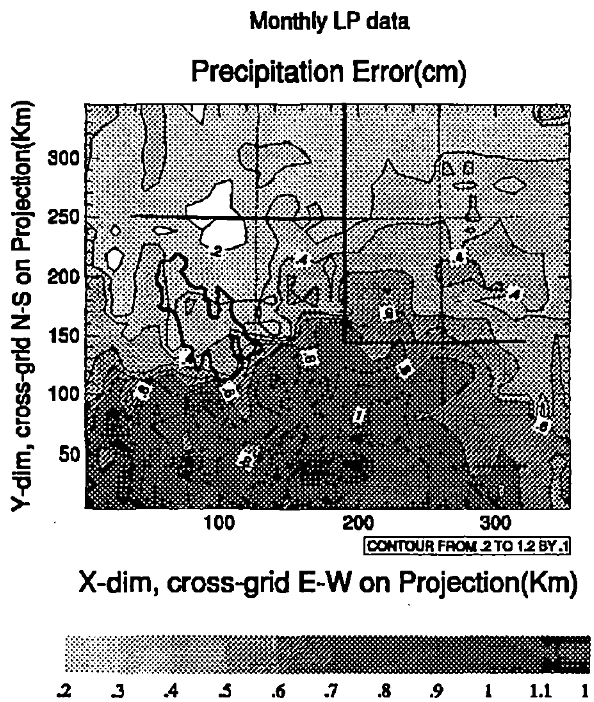


Figure 4. Observed precipitation error over model domain for September 1981. Observed errors are calculated using the local polynomial regression technique.

domain, it is easier to grid and the errors are smaller than daily precipitation. However, in this case the number of grid-points (1260) is an order of magnitude greater than the number of station observations (140), and thus we must use the resulting data set with caution. Figure 5 shows the location of the station observations and the grid-points. As shown, most grid-points do not have data within them or even in grid-points nearby. As such, the error estimates are much too low since they do not take this into account. Further experimentation with the LP found that a small change in the minimum number of points used in the regression allowed a specific grid-point in the Uinta Mountains to reach unrealistic values (120 mm). This points out the difficulty with extrapolating a high density gridded data set from a low density observation network. However, as this is the best that can be done, we will use this data set for comparision to the model experiments.

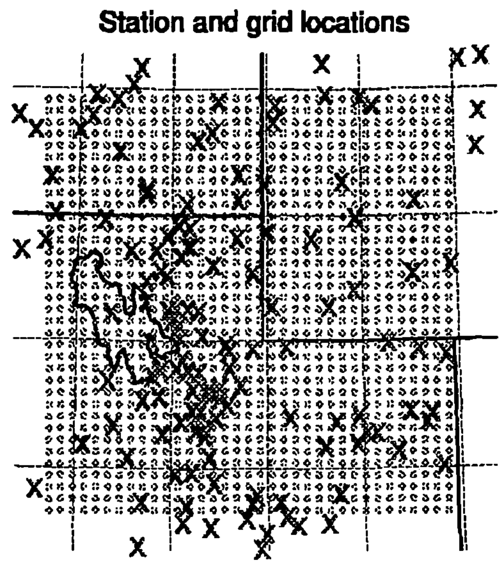


Figure 5. Locations of the observed stations and model grid points for the 10 km domain. Grid points are the circles and the stations are the x's.

Figure 6 shows the RegCM2 modeled precipitation for the same period as the observed precipitation shown in Figure 3. Notice that the precipitation is too highly concentrated on the mountain up-slopes and much too low over the entire grid domain. Also the simulation has five epicenters of tightly compact precipitation, and the observed is much more diffuse. However, the fact that the observed is diffuse probably

has more to do with the limited amount of observed data stations in the region. This behavior is similar to the 10 km simulation of *Nickerson et al.* [1986], and the 12.5 km simulation of *Zhang et al.* [1988]. The Nickerson simulation, was only a one-day simulation and they theorized that problems with the rainfall distribution had to do with spin-up. Zhang theorized that a 12.5 km simulation was too high a resolution for the hydrostatic model, and wondered if the vertical velocities were too low [*Zhang et al.*, 1988]. Our simulation was long enough to eliminate concerns with spin-up and

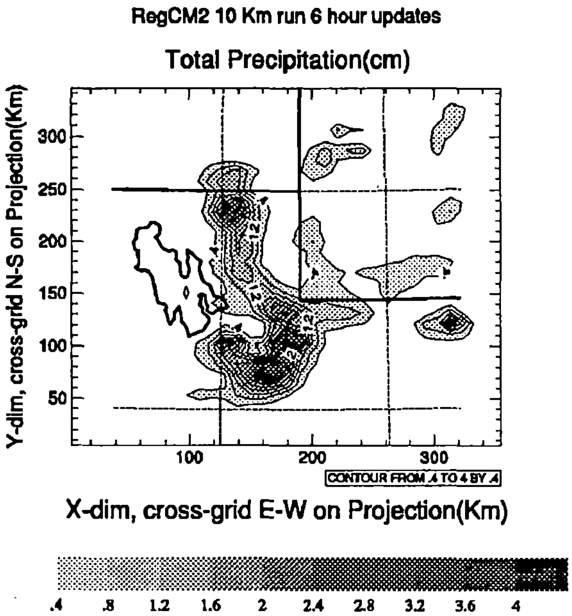


Figure 6. Accumulated precipitation from the RegCM2 model control simulation for September 1981.

therefore we theorize that problems with rainfall distribution are not related to this cause. Simulations over the complex terrain of Japan by *Kato* [1997] showed that while a 10 km hydrostatic simulation exaggerated maximum upward velocities, the impact on precipitation was very small. The net effect was an increase in precipitation amounts by 8%, and in area by 12%. Thus we conclude that the use of a non-hydrostatic model, compared to the use of our hydrostatic model, would not impact results significantly.

Let us investigate the possible causes of this increased distribution of precipitation over mountain up-slopes. As pointed out by *Sass and Christensen* [1995], the quality of a LAM simulation is dependent upon three factors – the quality of the LAM itself, the quality of the boundary relaxation scheme, and the quality of the boundary fields. We will investigate four issues: issues regarding the explicit cloud microphysics and cumulus parameterization, the need for cloud and rain water to be included on the LBCs, the quality of the input LBCs, and finally the LBC update frequency. To establish these points, we first examine theoretical issues associated with resolution and modeling of convective systems. Next, we assess the quality of the LBC data input into our model simulations. Lastly we examine some theoretical implications of the update frequency for LBC applied on a LAM. Finally, we will verify the conclusions of the theoretical analysis with the results of our model simulations.

Issues with High Resolution and Modeling of Cumulus Systems

As pointed out by *Molinari and Dudek* [1992], there are difficulties associated with mesoscale modeling at resolutions between 3 km and 25 km. Convective systems, though small in scale, have an important impact on the large-scale heat and moisture budget. At these resolutions there is not the clear separation of scale for cumulus parameterization assumed in its derivation. As reiterated by *Molinari and Dudek* [1992], *Cotton and Anthes* [1989] note that sub-grid-scale convective parameterizations become "muddy" and "not well posed" below 50 km resolution. Closure assumptions are also problematic since there is no observational evidence at these very fine resolutions. Even

the observational evidence at GCM scales is limited; but, at the 10 km scale there is not any observational support at all. Another problem they bring out is how to simulate mesoscale organized convection, mesoscale convective complexes (MCC), where individual cumulonimbus clouds organize to develop a stratiform storm system. For large-scales, the MCC will remain sub-grid-scale; but for high resolutions, a MCC represents a transition from the cumulus parameterization to grid-scale resolvable processes.

The next problem we will examine relates to the vertical velocities. The 10 km simulation will resolve some motions convectively parameterized at 50 km resolution. The vertical velocities will be higher and precipitation that is convectively parameterized at the lower resolution becomes grid-scale resolvable precipitation. However, since 10 km is still a large grid size, it is likely that the vertical velocities will not be high enough to produce the same amount of precipitation convectively parameterized at the lower resolution. Nor are the vertical velocities likely to be similar to that observed in cumulonimbus clouds. But, the model may produce enough grid-scale resolvable motion to stabilize the column [*Zhang et al.*, 1988]. Zhang recommends explicit cloud microphysics, and *Fritsch and Chappell* [1980] cumulus closure. In our initial experimentation with the model, we found that when the model was run with the implicit precipitation option – where any super-saturation immediately drops out as precipitation – and cumulus parameterization, the precipitation amounts were much too high to be realistic. By using explicit microphysics, we have additional loss terms, clouds form, and cloud and rain water evaporates. Experimentation also showed that the model seemed to be insensitive to the type of closure chosen for the cumulus parameterization. Although we do not examine the fraction of precipitation that comes from the cumulus parameterization, we can only speculate that it is similar to our current simulations, where it is around 10%. We also speculate that the reason precipitation was overpredicted in the implicit case was the lack of loss terms in the explicit formulation. We also speculate that changes to the explicit cloud microphysics will likely allow reasonable precipitation rates for the 10 km simulation.

Quality of the Input LBC Data to the 10 km RegCM2 Simulations

To establish the quality of our input LBC data, we compare these data to three data sets: a gridded surface temperature and precipitation data set, a rawinsonde profile, and the European Centre for Medium range Weather Forcasting (ECMWF) analysis. We compare the ECMWF analysis (interpolated to the 10 km grid) to the surface data set and rawinsonde. This validates the use of the ECMWF data set as a verification data set to use on the simulation input data.

ECMWF Data set Compared to Observed

The ECMWF data set is first interpolated to the 10 km grid domain. Next we find the grid-point in the 10 km grid domain nearest to the Salt Lake City rawinsonde, and compare the two. The ECMWF point nearest to the Salt Lake City (SLC) rawinsonde is within 2 degrees Celsius in temperature, 3 m/sec in winds, 3.0 g/kg in specific humidity, and within 20% in relative humidity. See **Figure 7** for a comparison of the ECMWF data-point to the SLC rawinsonde. Most of the differences are within the variability of the observed data set. However, the ECMWF is significantly too moist near the surface, and the variabilities in temperature and wind speed

are too low near the surface. The variability in relative humidity is also too low.

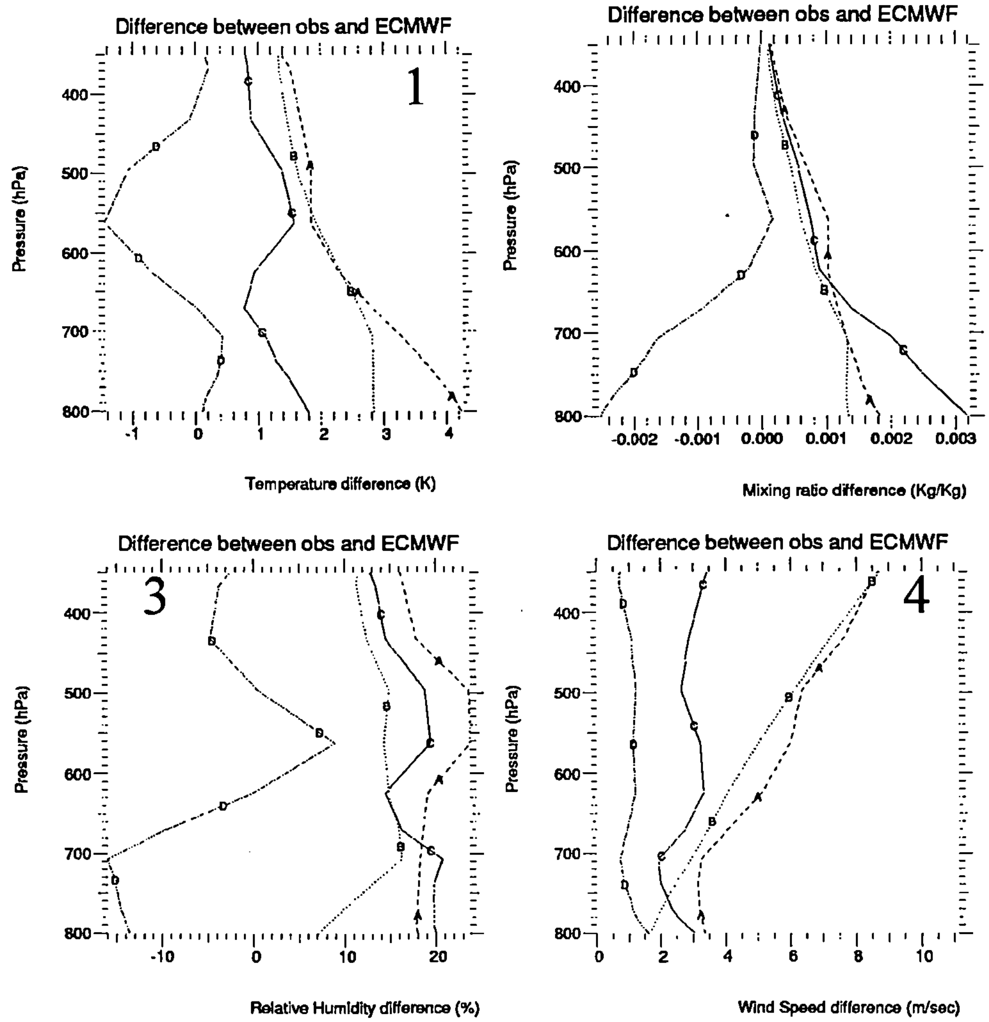


Figure 7. (1) Average and RMS differences (SLC - ECMWF) from ECMWF dataset to SLC rawinsonde, as well as standard deviations for temperature. The ECMWF data was first interpolated to the 10km model domain and then the grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B" ECMWF dataset standard deviation. "C", RMS difference between ECMWF and SLC. "D", average difference between ECMWF and SLC. (2) Average and RMS differences (SLC - ECMWF) from ECMWF dataset to SLC rawinsonde, as well as standard deviations for water vapor mixing ratio. The ECMWF data were first interpolated to the 10km model domain and then the grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B" ECMWF dataset standard deviation. "C", RMS difference between ECMWF and SLC. "D", average difference between ECMWF and SLC. (3) Average and RMS differences (SLC - ECMWF) from ECMWF dataset to SLC rawinsonde, as well as standard deviations for water vapor mixing ratio. The ECMWF data were first interpolated to the 10km model domain and then the grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B". ECMWF dataset standard deviation. "C", RMS difference between ECMWF and SLC. "D", average difference between ECMWF and SLC. (4)Average and RMS differences (SLC - ECMWF) from ECMWF dataset to SLC rawinsonde, as well as standard deviations for relative humidity. The ECMWF data were first interpolated to the 10km model domain and then the grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B" ECMWF dataset standard deviation. "C", RMS difference between ECMWF and SLC. "D", average difference between ECMWF and SLC.

Simulation Input Data set Compared to Observed

The 10 km input data set can be compared to observed in two ways: examining the profile by comparing one model point to the rawinsonde, or examining the surface temperature by comparing all model points to a gridded observed data set. We will do both. In Figure 8, we compare the point in the input data set closest to SLC to the SLC rawinsonde.

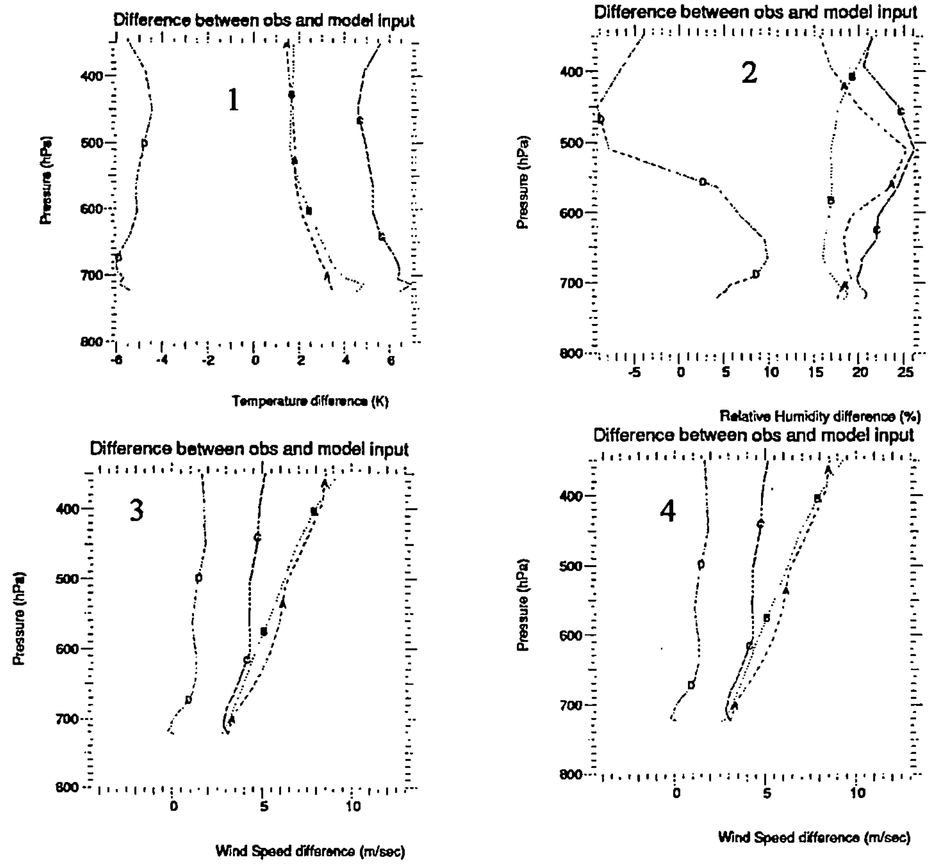


Figure 8. **(1)** Average and RMS differences (SLC - input) from input dataset to SLC rawinsonde, as well as standard deviations for temperature. The grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and SLC. "D", average difference between input and SLC. **(2)** Average and RMS differences (SLC - input) from input dataset to SLC rawinsonde, as well as standard deviations for water vapor mixing ratio. The grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B", input dataset standard deviation. "C", RMS difference between input and SLC. "D", average difference between input and SLC. **(3)** Average and RMS differences (SLC - input) from input dataset to SLC rawinsonde, as well as standard deviations for relative humidity. The grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and SLC. "D", average difference between input and SLC. **(4)** Average and RMS differences (SLC - input) from the input dataset to SLC rawinsonde, as well as standard deviations for wind speed. The grid point closest to SLC was used to compare to the rawinsonde. "A", Salt Lake City rawinsonde dataset standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and SLC. "D", average difference between input and SLC.

The input data set's temperature is 6 degrees Celsius too warm, and up to 2 g/kg too moist. Winds are within 4 m/sec of observed. Relative humidity is within about 20% of observed, being 10% too low near the surface and 10% too high around 500

hPa. Both temperature and specific humidity are too high, but the increase compensates for each other to give good agreement with relative humidity. The differences in humidity and wind speed are at the observed variability, but temperature is significantly too warm.

The RMS difference between the observed monthly surface temperatures over the domain is 1.3 degrees Celsius. In Figure 9 we compare the RMS differences between the input 10 km data set of surface

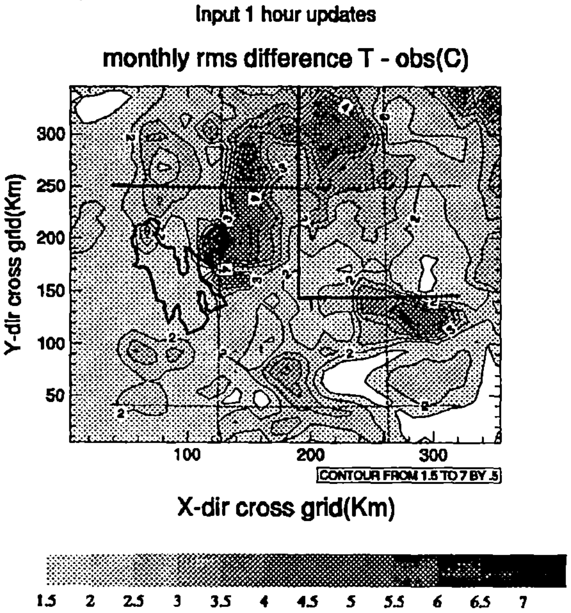


Figure 9 Contour plot of the RMS difference between input dataset and observed daily T_{ave} for September, 1981. Observed temperature calculated using the local polynomial regression technique on the mean of daily maximum and minimum temperatures. Input dataset temperature is calculated by finding the mean of the 6 hour data.

temperature and the gridded observed surface temperature. The observed temperature data set is created with observed station temperatures using the LP regression and the error estimates are given in Figure 10. RMS difference errors are roughly 1 degree Celsius higher than the observed LP error estimates. Thus we conclude there is reasonable agreement between the observed and the 50 km input data set, with the exception that the input data set is much too moist and warm near SLC.

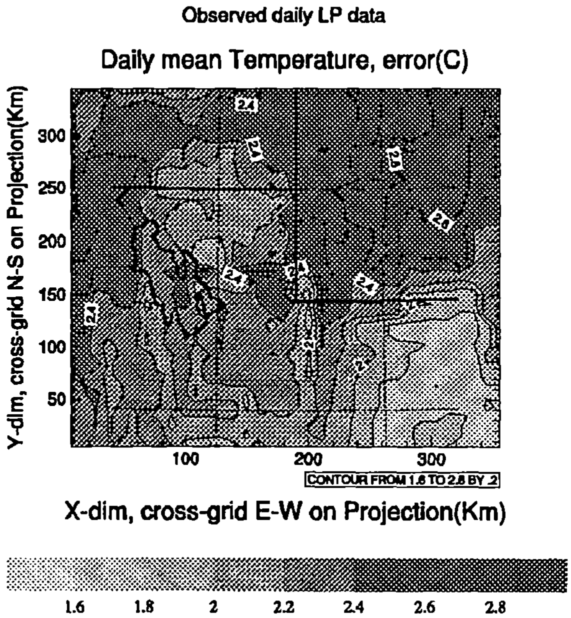


Figure 10 Average observed temperature error for September 1981. Observed error calculated using the local polynomial regression technique.

Simulation Input Data set Compared to the ECMWF Data set.

The ECMWF data set, just validated in section 0, is now compared to the 10 km simulation input data set. As shown in Figure 11, the 50 km data set is moister than ECMWF. The input data set's temperature is within 2 degrees of ECMWF and winds are too fast by 4 m/sec. Specific humidity is much too high by 2 mg/kg, with significant differences throughout the entire profile. Relative humidity is also conspicuously higher than observed. The differences in temperature and wind speed are within observed variability. However, humidity is much moister than ECMWF. Since the ECMWF data set itself is too moist near SLC, this certainly indicates the input data set is much moister than reality. The input data set was also shown to be significantly

too warm near SLC, which means considerable variation can exist between individual data points.

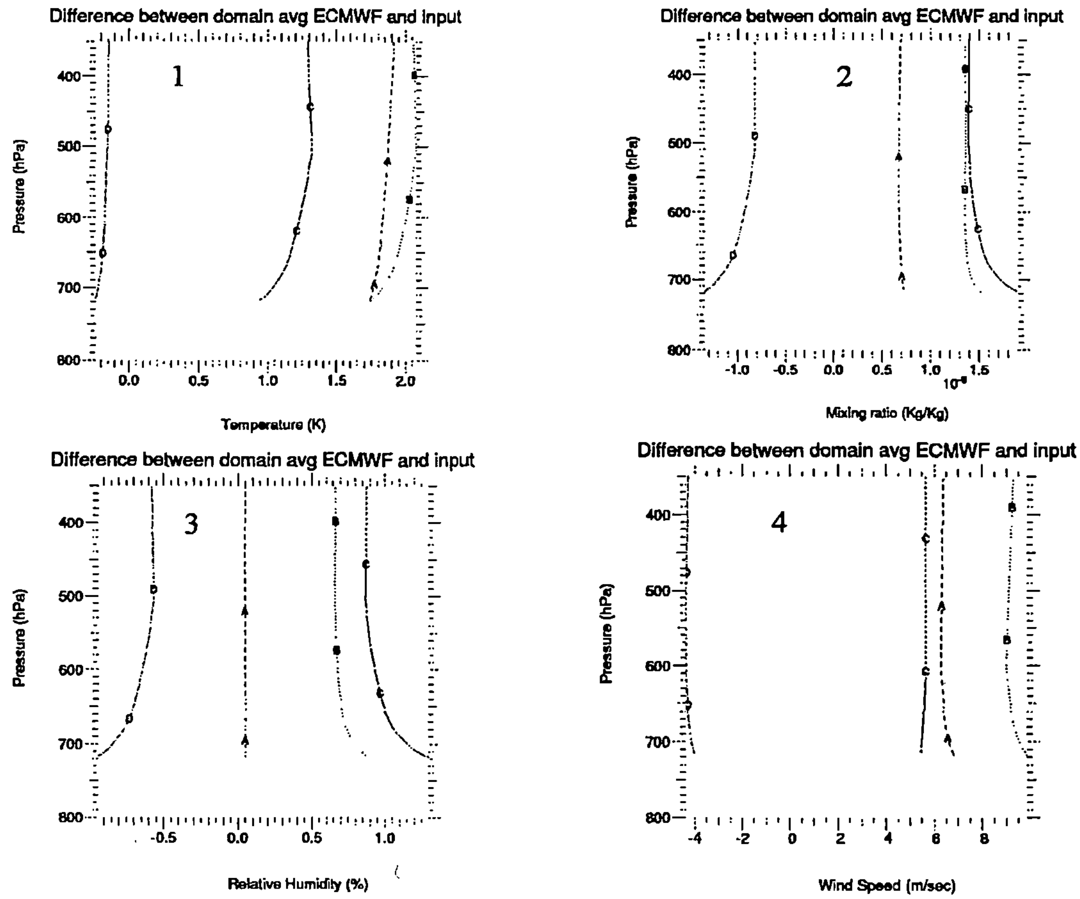


Figure 11. (1) Average and RMS differences (ECMWF - input) of domain averaged temperature between the ECMWF and input datasets, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. Only datapoints for which the elevation of the ECMWF dataset is within 3 m of the input dataset are considered in the domain average. "A", ECMWF standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and ECMWF. "D", average difference between input and ECMWF. (2) Average and RMS differences (ECMWF - input) of domain averaged water vapor mixing ratio between the ECMWF and input datasets, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. Only datapoints for which the elevation of the ECMWF dataset is within 3 m of the input dataset are considered in the domain average. "A", ECMWF standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and ECMWF. "D", average difference between input and ECMWF. (3) Average and RMS differences (ECMWF - input) of domain averaged relative humidity between the ECMWF and input datasets, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. Only datapoints for which the elevation of the ECMWF dataset is within 3 m of the input dataset are considered in the domain average. "A", ECMWF standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and ECMWF. "D", average difference between input and ECMWF. (4) Average and RMS differences (ECMWF - input) of domain averaged wind speed between the ECMWF and input datasets, as well as standard deviations. Data is averaged over each sigma layer and the average pressure for that sigma layer given. Only datapoints for which the elevation of the ECMWF dataset are within 3 m of the input dataset are considered in the domain average. "A", ECMWF standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and ECMWF. "D", average difference between input and ECMWF.

In conclusion, we find that the errors in our input data set are similar to the errors in both the observed data set and the ECMWF data set. Thus we surmise that our input data set should provide reasonable LBC for the high-resolution simulation. The one concern that remains is that the input data set is certainly much more moist than that observed.

Theoretical Mechanism for Artificially Enhanced LBC Storms

With 6-hour updates starting at 12:00 am Greenwich Mean Time (GMT), we miss some of the daily extreme events. Accounting for daylight saving time, updates at 12:00 am, 6:00 am, 12:00 pm, and 6:00 pm GMT are updates at 5:00 pm, 11:00 pm, 5:00 am, and 11:00 am local standard time. Thus the linear time interpolation will smooth over the hottest part of the day. And, the LBCs between 11:00 am and 5:00 pm will be too cool, with a maximum difference at about 2:00 pm. Since the interior of the model properly responds to solar heating, the difference in temperature on the interior to the LBC produces a temperature gradient near the LBC. This temperature gradient, if significant, can cause cool air to be artificially advected in from the lateral boundaries. As cool air is advected into the system, it can cause enough cooling to bring the air to saturation and create an artificially enhanced storm event. The reason this does not happen in regional models with larger domains is that the temperature gradient is applied over a larger area. The gradient is applied over the "relaxation" region on the lateral boundaries where the adjustment scheme blends input and model data. For our 50 km run the relaxation region consists of the outer 11 boundary grid-points. This implies that an error in the outer boundary has to be applied over a distance of 550 km (11 cells of 50 km each) before the error changes the gradients near the boundary.

An error of 3 degrees Celsius on the boundary produces an implied gradient of $5.5\text{x}10^{-6}$ C/m. For the 10 km grid the relaxation region is only five cells or 50 km. The magnitude of the error in the 10 km run is an order of magnitude larger than the implied error in the 50 km run.

This same temperature error of 3 degrees Celsius in the LBC now produces an implied gradient of $6\text{x}10^{-5}$ C/m. The model will apply this gradient both in horizontal diffusion and advection, but for comparison purposes we will examine horizontal advection only.

Advection is a driving term in the continuity, energy, and momentum equations. It is the gradient in wind speed and the given quantity examined (such as temperature). Let us calculate the advection of temperature for a simple case. Let us assume a constant wind of 5 m/sec and an implied temperature gradient of 3 degrees Celsius applied over the boundary region. Consequentially, the implied advective cooling will be $-3\text{x}10^{-4}$ C/sec. This is of the same magnitude as the dominant terms in the energy equation. For the 50 km simulation this term will be less than one tenth that size and therefore it will not be a dominant term. Also, since a significant part of the LBC for the 50 km domain is over ocean, an error of 3 degrees Celsius is less likely than in the 10 km domain where an error of that magnitude is very common. This is because oceans experience an insignificant diurnal variation compared to landmasses. Thus errors in time interpolations are much smaller over the ocean.

In equations 1 and 2, we show the formula for water vapor saturation. When we use the values for 700 hPa and 280 K we get a saturation-mixing ratio of 0.90 g/kg. With a

3-degree cooling to 277 K, we get saturation at 0.73 g/kg. Thus the 3-degree error artificially changes 80% relative humidity to 100% relative humidity.
The formula for water vapor saturation is:

$$q_{SAT} = 0.622 p_{SAT} / (p - 0.378 p_{SAT}) \quad (1)$$

$$p_{SAT} = 6.11 \exp(19.84659 - 5418.12/T) \quad (2)$$

where p_{SAT} in equation 2 is the saturation pressure (hPa); p is the pressure (hPa); q_{SAT} is the water vapor mixing ratio at saturation (kg/kg); and T is the temperature (K).
It is quite plausible that as this sort of error gets advected through the system, it can artificially enhance a storm moving through the domain. As we see in the results section, spurious storms do not seem to be created through this mechanism, but they can alter a storm system as it comes into the model domain.

Results

In the discussion we establish that the input data set to our 10 km simulation compared well to the observed data. Although there are some concerns with the data set (see section 0), the input fields are near to the observed values and should provide good input LBCs. In section 1 we introduced two problems with our 10 km simulation. The first was the general lack of precipitation over the entire grid-cell, and the second was the overabundance of precipitation associated with mountain up-slopes.
We believe the first problem (lack of precipitation) is due to two problems. The first and primary problem has to do with the cloud microphysics. The explicit cloud parameterization creates thick clouds that do not produce significant precipitation. The cloud parameterization (explicitly the auto-conversion formulation) needs to be changed so that more of the cloud water is converted into precipitation. The second problem has to do with our lack of providing cloud and rain water as input at the LBC. This causes two changes. One, precipitation will fall near the lateral boundaries, a region currently almost completely dry. Second, clouds will be "spun-up" to precipitating levels faster, since they enter the model domain as cloud rather than having to develop within the small model domain.

The second problem (abnormally high amounts of precipiation on up-slopes) was due to the "artificially enhanced LBC storms" that were discussed in section 0. A closer examination of these two problems is warranted. Let us first turn our attention to the general lack of precipitation in the model domain. In Figure 12 we compare the total precipitable water of the input data set to ECMWF over the 10 km simulation domain.

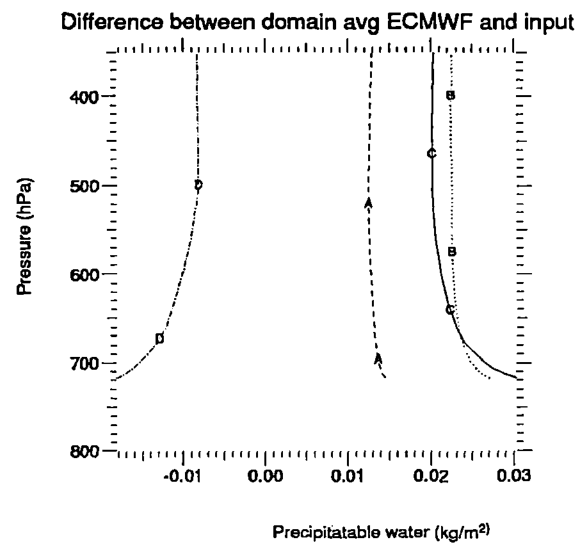


Figure 12. Average and RMS differences (ECMWF - input) of domain averaged total water mass between the ECMWF and input datasets, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. Only datapoints for which the elevation of the ECMWF dataset is within 3 m of the input dataset are considered in the domain average. "A", ECMWF standard deviation. "B" input dataset standard deviation. "C", RMS difference between input and ECMWF. "D", average difference between input and ECMWF.

Precipitable water is the total water mass in the atmosphere. Total mass over the domain of the observed precipitation is 3.08×10^{12} kg while the same value for the 50 km and 10 km simulations are 4.24×10^{12} kg and 2.09×10^{11} kg, respectively. Thus total precipitable water is an order of magnitude higher than the precipitation. Likewise, cloud water is an order of magnitude higher than precipitation as shown in Figure 13.

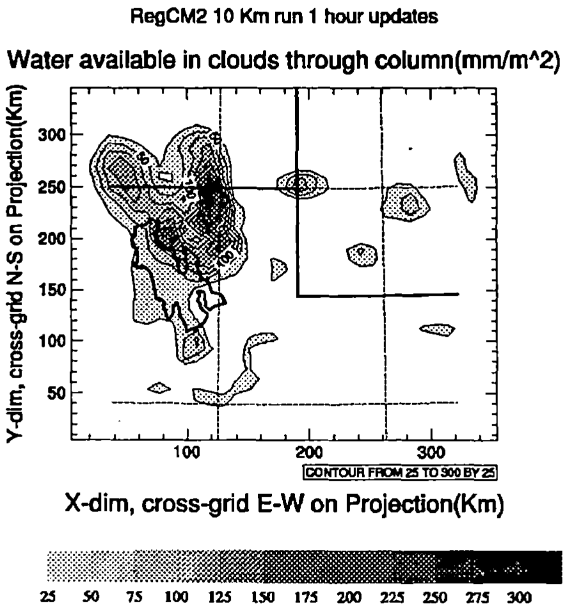


Figure 13. Control simulation total column cloud water totaled over September 1981 in mm of water per unit area.

The input 50 km simulation is a little high, as we can see from these values. However, the resulting 10 km simulation is an order of magnitude lower than the observed and 50 km simulation. See Figure 14 for a plot of the average precipitation over the domain area for each day. The input precipitable water as shown in Figure 12 is slightly high, which confirms the excess precipitation in the 50 km simulation, but disagrees with the low level of precipitation in the 10 km simulation. We must consider the fact that the same conditions appear for the 50 km simulation as well. Values that are interpolated from the 50 km simulation drive the 10 km simulation. So why does the 10 km simulation receive so much less precipitation than the 50 km simulation if the two have similar conditions?

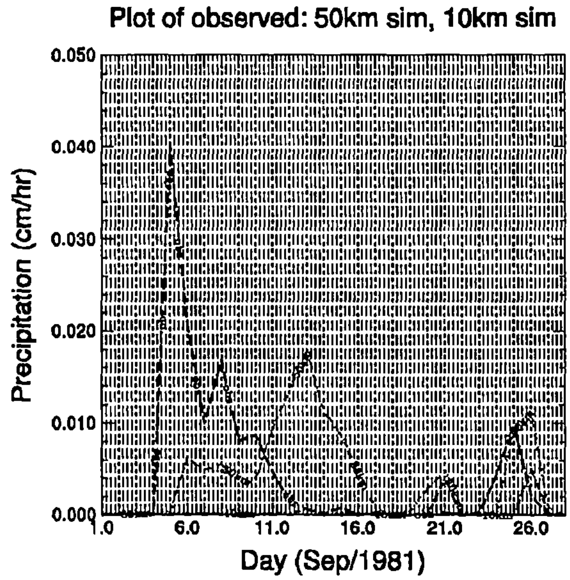


Figure 14. Comparison of domain averaged precipitation with time. Observed (black), average of 50 km simulation over the region (green), and average of the 10 km simulation (red). Dotted grid marks each day, and the dashed grid marks each 6 hour interval.

The bias in precipitable water leads us to consider possible problems with the interpolation of the 50 km simulation for input into the 10 km simulation. A concern with the interpolation of the 50 km data set is that the interpolation was done on sigma levels rather than pressure surfaces. Interpolation on sigma surfaces mixes the horizontal and vertical derivatives and impacts the end result. Extrapolating temperatures for 10 km terrain beneath the 50 km surface is also problematic. *Trenberth et al.* [1993] dealt with these issues in detail. We implement their procedures for vertical interpolation. This amounts to interpolating the 50 km surface, first to pressure levels, then horizontally to the 10 km grid; and, at last, vertically interpolating the pressure levels to the final 10 km sigma coordinates. We also used the *Trenberth et al.* [1993]

procedures to extrapolate temperatures below ground. Our treatment of humidity also differs. Because water vapor mixing ratio is conserved in the boundary layer, we interpolate on the mixing ratio near the surface rather than relative humidity. At higher levels we interpolate on the relative humidity as in *Trenberth et al.* [1993]. We also use a simple log profile to compensate for the winds that became closer to the ground after the interpolation. Figure 15 shows the difference between our new LBC data set and the original control simulation. The differences are negligible. The total water mass is unchanged between the two data sets. The differences in other variables are very small and much less than the variability in the data sets.
There is another source in the water budget that we have not considered. That source is cloud and rain water.

The 50 km simulation handles precipitation in a different manner than the 10 km simulation. The 50 km simulation uses implicit cloud physics, meaning that cloud water is diagnosed rather than prognostically solved for. The 10 km simulation uses explicitly prognosed cloud and rain water. Since the 50 km simulation did not have cloud and rain water, we did not include it as part of the LBCs. Physically this implies that no cloud or rain water gets advected into the system from outside the system. Anyone familiar with a desert area realizes that the majority of the water comes from storms that have developed outside the area. The implicit form of precipitation in the 50 km simulation means that precipitation immediately appears on the surface as rain. Thus, there is no time for this rain water to be advected into the 10 km domain. Were the interface between the 50 km and 10 km simulations realistic, the cloud and rain water from the 50 km domain would be advected into the 10 km domain. Since the 50 km simulation was done without cloud and rain water, we can only examine the total amount of water that would be advected into the system from the 50 km domain. To estimate this we total the amount of rain over the 50 km cells on the 10 km LBC. To make sure that we only get the rain that would be advected into the system, we only sum those grid-points that have a wind moving into the 10 km domain. We estimate that up to 1×10^{12} kg of rain water would be available for advection into the 10 km simulation in this manner. This raises the precipitation bias to about half of the observed. Most of this water will be deposited near the lateral boundaries, and not impact the total simulation. However, some portion will enter the system as cloud water and affect the interior solution as the source is driven into the simulation domain. Since this source of water only accounts for a fraction of the precipitation deficit, there must be other sources of water to consider.

Two other sources of additional water are the cloud water in the system, and water that can be released from the cumulus parameterization. We will consider cloud water first. The total column cloud water totaled over the domain is 74 mm/m^2. As such, there is sufficient water that could be converted into precipitation while still allowing clouds in the system. See Figure 13 for the distribution of total column cloud water for the entire period through the model domain. Notice that cloud water is fairly evenly distributed, so converting more cloud water into rain water will create a water source evenly distributed over the domain. Since cloud water forms as water vapor is saturated, it is unlikely that the precipitation deficit is because of supersaturation of water vapor. However, the cloud water physics could be tuned to allow for more precipitation.

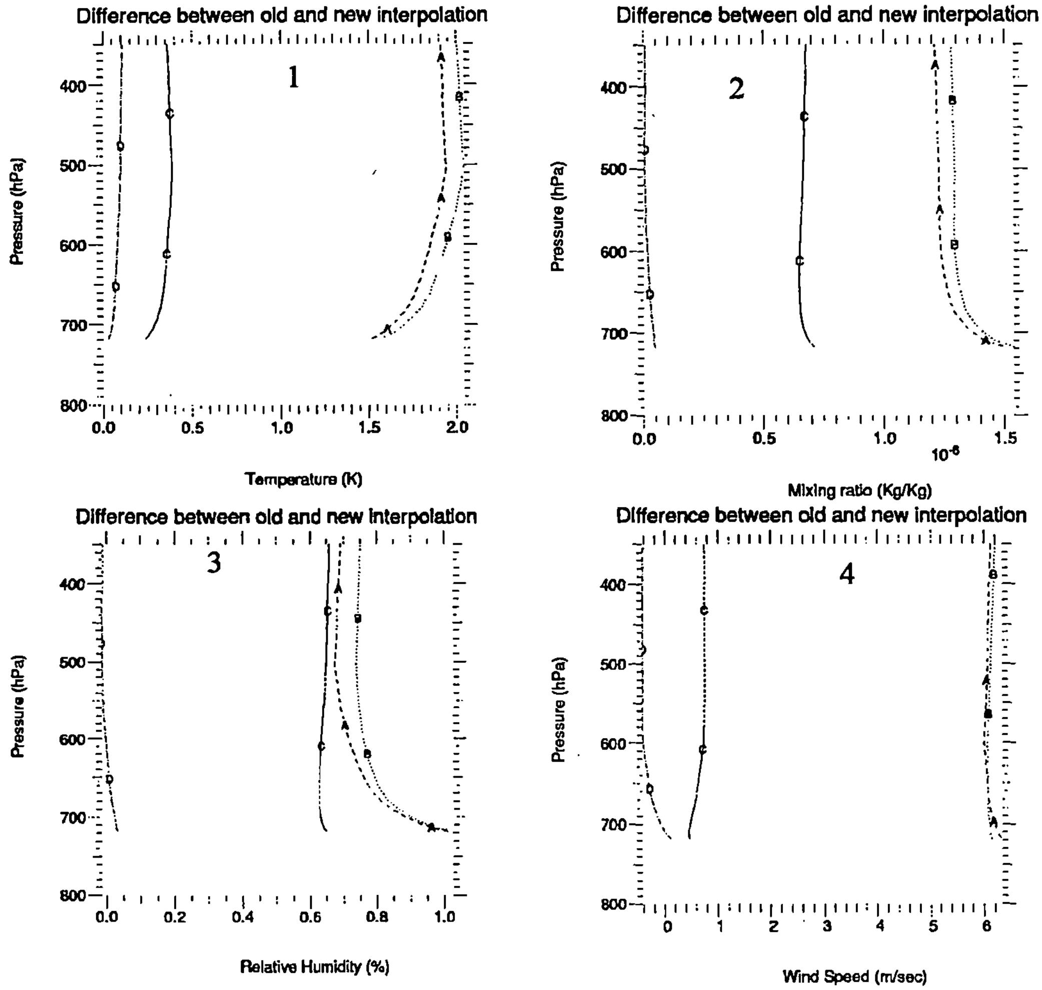


Figure 15 (1) Average and RMS differences (new - old) between the domain average for the control input dataset and the new interpolation methods for temperature, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. "A", "new" standard deviation. "B" "old" dataset standard deviation. "C", RMS difference between "new" and "old". "D", average difference between "new" and "old".

(2) Average and RMS differences (new - old) between the domain average for the control input dataset and the new interpolation methods for water vapor mixing ratio, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. "A", "new" standard deviation. "B" "old" dataset standard deviation. "C", RMS difference between "new" and "old". "D", average difference between "new" and "old".

(3) Average and RMS differences (new - old) between the domain average for the control input dataset and the new interpolation methods for relative humidity, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. "A", "new" standard deviation. "B" "old" dataset standard deviation. "C", RMS difference between "new" and "old". "D", average difference between "new" and "old".

(4) Average and RMS differences (new - old) between the domain average for the control input dataset and the new interpolation methods for wind speed, as well as standard deviations. Data are averaged over each sigma layer and the average pressure for that sigma layer given. "A", "new" standard deviation. "B" "old" dataset standard deviation. "C", RMS difference between "new" and "old". "D", average difference between "new" and "old".

The most important term is the auto-conversion term that initiates the creation of rain droplets. RegCM2 uses a *Kessler* [1969] expression for auto-conversion as follows:

$$P_{RC} = k_1(q_c - q_{cth}) \qquad (3)$$

where $q_{cth} = 0.41$ g/kg, $k_1 = 1\times10^{-3}$ 1/m, q_c is the cloud water mixing ratio, and P_{RC} acts as a source term for rain water and a loss term for cloud water.

Anthes et al. [1987] mention that this parameterization is sensitive to resolution, and *Hsie* [1983] did some limited sensitivity studies. Thus, it is not surprising that some tuning may be in order for a different resolution. There is a substantial amount of cloud water above the threshold of 0.3 g/kg (from Figure 10). Cloud water is also spread out over the entire domain area, which should improve the grid-averaged statistics. Lowering the auto-conversion threshold will increase precipitation. In experiments at 20 km resolution over a similar domain for July 1981, we found lowering the threshold from 0.5 g/kg to zero doubled the maximum precipitation amounts. Assuming total precipitation doubles, this still brings us quite short of the precipitation deficit for our situation. Although sufficient water is available in the clouds to account for the precipitation deficit, realistic changes to the auto-conversion term do not allow large enough increases in precipitation. Lowering the auto-conversion threshold is akin to allowing precipitation to form in clouds less "thick" with moisture. As we discussed in section 2.3, the model is at an awkward resolution. The resolution is high enough that most of the convective activity is resolved. However, the grid size is large enough that the vertical velocities are likely to be much lower than might be observed in an actual cumulonimbus cloud system. Therefore, the model does not produce as much condensation as might have been produced in the cumulus parameterization. Hence, the condensation that it does produce must be appropriately allocated. Allowing clouds to form at saturation levels less than 100% may be required to account for the precipitation deficit.

The next water source to consider is the possibility of getting more precipitation out of the cumulus parameterization. However, first we consider this likelihood. The cumulus parameterization accounted for approximately 35% of the precipitation for the 50 km simulation over this region for September 1981. At a higher resolution, we expect that much of the activity dealt with in the larger-scale simulation will be resolved at the 10 km level; and we expect a reduction in the percentage of precipitation coming from a subscale source. Even the same percentage as the 50 km simulation (35%) would not explain the precipitation deficit. What was observed was approximately 0.5% of the amount in the 50 km simulation. Increasing this percentage to even 5% may take substantial changes to the parameterization. In the end, these changes seem unlikely to affect the overall quality of the model. Thus, it seems the best candidate for tuning changes is the auto-conversion parameterization. One possibility is to replace the auto-conversion parameterization with a more modern expression that is less sensitive to changes in resolution. We may also need to allow cloud water to form at subsaturated levels.

The second problem we wish to discuss is artificially enhanced LBC storms. As a first check for LBC storms, we run the model at a higher LBC update and observe if the simulation changes significantly. If the update frequency is sufficiently high, we expect almost no difference between the two

solutions. In Figure 16 we show a simulation with hourly LBC updates. The precipitation distribution has changed, but the amounts

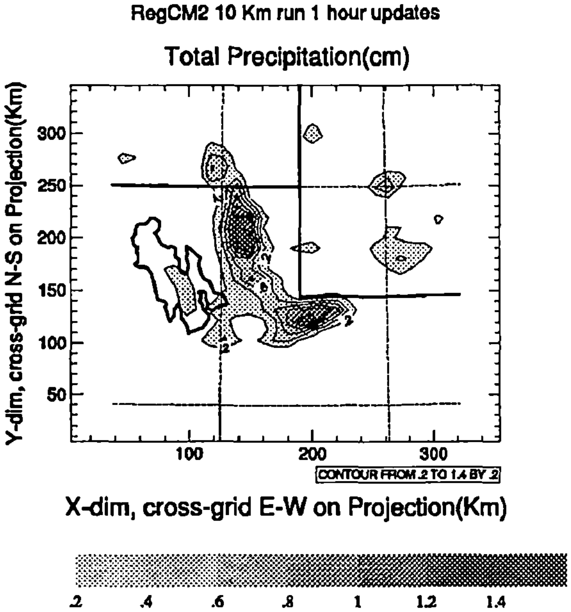


Figure 16. Precipitation for 10 km simulation with hourly LBC updates.

are much less than the control simulation with 6 hour LBC updates. We need to examine the simulation in more detail to understand how and why the simulation changed. Figure 17 shows the difference between the grid-scale-averaged precipitation for the control to the 1-hour update. There is a certain amount of low-level noise that is expected, but there are seven spikes of significant magnitude in the series. Of the seven, six show increased precipitation in the control run and only one shows decreased precipitation. Most of these spikes also initiate at the 6-hour period between 11:00 am and 5:00 pm, and peak this same time as well. So although this evidence is circumstantial, it seems to indicate artificial enhancements of several of these storms. However, at the same time any statistical measure of performance that we apply to either simulation, specifically the RMS differences of precipitation with observed and zero level threshold threat scores, shows that neither simulation is preferred. But, the reason for this is the model's unreasonable performance at this resolution and entirely too low precipitation amounts. Once the changes to the explicit cloud water formulation are made, this experiment needs to be performed again.

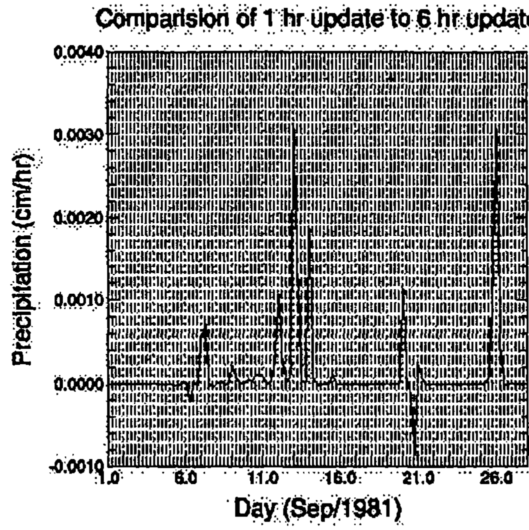


Figure 17. Comparison of grid-scale average precipitation from control to 1 hour update case. Dotted grid marks each day, and the dashed grid marks each 6 hour interval.

Discussion and Conclusions

We have discussed some of the difficulties associated with modeling precipitation in LAM at the 10 km scale. Our simulations exhibited the same behavior shown by other simulations – not enough precipitation over the grid domain coupled with too much precipitation over the mountain up-slopes. We have done extensive analysis on these simulations and concluded the following. Cloud and rain water inputs are important to include on the LBC. So LBC update frequencies and total domain area impact the final solution.

When we examined the total water budget, we found the precipitable water to be too low in the 10 km system. However, as the same bias was displayed in the 50 km simulation, we felt the answer will be found elsewhere. One bias is the lack of input

cloud and rain water into the system. This bias may account for 50% of the total deficit in precipitation. We believe the rest of the bias has to do with the error in precipitable water already discussed, as well as the lower than observed variability in relative humidity. At this point some tuning to the explicit cloud and rain water physics can be done. Although initially we did preliminary experimentation with the cloud water physics, our results were unsuccessful. The total bias was too large for a single change to one parameter to affect the results. By doing extensive analysis and attacking the problem on several fronts, we think we have arrived at a successful solution to the problem.

Now we consider concentration of rainfall on windward up-slopes. Much of this problem has to do with the grid-scale average problem. Since precipitation does not occur en masse over the grid domain, the only precipitation that does form is via orographic uplift. By allowing cloud and rain water to be advected in from the LBC, more precipitation will form over the entire grid domain.

As to the point brought up by *Kato* [1997] in regard to "determining the optimal dynamical framework," our contribution seems to be increased understanding of one method – one-way nesting. There are some problems unique to this setup that are exasperated at high resolutions. Sufficiency of domain size and LBC update frequency is an important issue for one-way nesting. One-way nesting is a popular method of incorporating large-scale information into a LAM. It is likely that it will remain so. Even models with two-way nesting require one-way nesting with large-scale input fields. As GCM resolutions increase, the technology to one-way nest LAMs to very high resolution GCM simulations will become important, hence, the utility of our study.

Kato [1997] continues with the question of "the optimal precipitating scheme." We found that explicit cloud and rain water were important components of these 10 km simulations. Our contribution is to point out the need for explicit precipitation physics in high-resolution LAM simulations. We also point out that for our domain, convective paramaterization did not play a role. As to the question of the utility of high resolution climate simulations – we suggest this issue warrants further study. Observation networks do not exist at resolutions required to establish the fine details of these simulations. And, there are significant tuning problems to be worked out with such simulations.

Finally, can LAMs be used to provide accurate answers to regional concerns of climate change and run impact models? Once the model is sufficiently tuned, we believe so. However, the one-way nesting to 50 km and 10 km simulations is limited. This is due to the fact that the biases of the GCM are input into the 50 km and 10 km models. This is compounded with the biases of the 50 km and 10 km models and affects the final 10 km simulation. This is another reason that accurate analysis of the model simulations must be done at each step. It is also important to understand the biases inherent in going between each model simulation, gaining an understanding of the total bias of the system. This is why it is vital to compare results for observed sequences, such as we have done here, in order to gain accurate understanding of the total error in the system.

References

Anthes, R. A., Hsie E.-Y., Kuo Y.-H., Description of the Penn State/NCAR mesoscale model version 4 (MM4), *NCAR Tech. Note, TN-282+STR*, 66 pp., National Center for Atmospheric Research, Boulder, Colo., USA, 1987.

Briegleb, B. P., Delta-Eddington approximation for solar radiation in the NCAR Community Climate Model, *J. Geophys. Res.*, *97*, 7603-7612, 1992.

Cotton W. R., and R. A. Anthes, *Storm and Cloud Dynamics*, 883 pp., Academic Press, San Diego, Calif., 1989.

Davies, H. C. and R. E. Turner, Updating prediction models by dynamical relaxation: An examination of the technique, *Q. J. R. Meteorol. Soc.*, *103*, 225-245, 1977.

Dickinson, R. E., A. Henderson-Sellers, P. J. Kennedy, Biosphere-atmosphere transfer scheme (BATS) version 1E as coupled to the NCAR community climate model, *NCAR/TN-387+STR*, 69 pp., National Center for Atmospheric Research, Boulder, Colo., USA, 1993.

Fritsch, J. M., and C. F. Chappell, Numerical prediction of convectively driven mesoscale pressure systems. Part I: Convective paramertization., *J. Atmos. Sci.*, *37*, 1722-1733, 1980.

Giorgi, F., Sensitivity of simulated summertime precipitation over the western United States to different physics parameterizations, *Bull. Am. Meteorol. Soc.*, *119*, 2870-2888, 1991.

Giorgi, F., Simulation of regional climate using a limited area model nested in a general circulation model, *J. Clim.*, *3*, 941-963, 1990.

Giorgi, F., and G. T. Bates, The climatological skill of a regional model over complex terrain, *Mon. Weather Rev.*, *117*, 2325-2347, 1989.

Giorgi, F., G. T. Bates, and S. J. Nieman, Simulation of the arid climate of the Southern Great Basin using a regional climate model. *Bull. Amer. Meteorol. Soc.*, *73*, 1807-1822, 1992.

Giorgi, F., G. T. Bates, and S. J. Nieman, The multiyear surface climatology of a regional atmospheric model over the western United States, *J. Clim.*, *6*, 76-95, 1993a.

Giorgi, F., M. R. Marinucci, and G. T. Bates, Development of a second-generation regional climate model (RegCM2). Part I: Boundary-layer and radiative transfer processes, *Mon. Weather Rev.*, *121*, 2794-2813, 1993b.

Grell, G. A., J. Dudhia, and D. R. Stauffer, A description of the fifth generation Penn State/NCAR mesoscale model (MM5), NCAR Tech. Note, NCAR/TN-398+STR, 121pp., National Center for Atmospheric Research, Boulder, Colo., USA, 1994.

Holtslag, A. A. M., E. I. F. de Bruijin, and H. L. Pan, A high-resolution airmass transformation model for short-range weather forecasting, *Mon. Weather Rev.*, *118*, 1561-1575, 1990.

Holtslag, A. A. M., and B. A. Boville, Local versus non local boundary-layer diffusion in a global climate model, *J. Clim.*, *6*, 1825-1842, 1992.

Hong, S.-Y., and H.-M. H. Juang, Orography blending in the lateral boundary of a regional model, *Mon. Weather Rev.*, *126*, 1714-1718, 1998.

Hostetler, S. W., and F. Giorgi, Use of a regional atmospheric model to simulate lake-atmosphere feedbacks associated with Pleistocene lakes Lahontan and Bonneville, *Clim. Dyn.*, *7*, 39-44, 1992.

Hsie, E.-Y., Frontogenesis in a moist atmosphere, Ph.D. Thesis, Department of Meteorology, Pennsylvania State University, Penn, 251 pp., 1983.

Hsie, E.-Y., and R. A. Anthes, Simulations of frontogenesis in a moist atmosphere using alternative parameterizations of condensation and precipitation. *J. Atmos. Sci.*, *41*, 2701-2716, 1984.

Kato, T., Hydrostatic and non-hydrostatic simulations of moist convection: review and further study, *Meteorol. Atmos. Phys.*, *63*, 39-51, 1997.

Kessler, E., III, On the distribution and continuity of water substance in atmospheric circulation's, *Meteorol. Monogr.*, *AMS*, *No. 27*, 84 pp., 1969.

Lall, U., and K. Bosworth, Nonparametric statistical inference and function estimation for hydrologic space time data, *Trends Hydrol.*, CSRI, 1-19, 1994.

Madala, R. V., Efficient time integration schemes for atmosphere and ocean models, in *Infinite-Difference Techniques for Vectorized Fluid Dynamics Calculations*, edited by D. L. Book, pp. 56-74, Springer-Verlag Publ., New York, 1981.

Matthews, D. M., F. Giorgi, and G. T. Bates, Regional model simulation of an extreme precipitation event in the Colorado River Basin, in *Proceedings of the Third Symposium on Global Change*, AMS, Boston, Mass., 20-23, 1992.

Medina, J. G., Application of a simple local-scale numerical model in the study of altered climate impacts on watershed precipitation, in *Proceedings of the Second Symposium on Global Change Studies*, AMS, New Orleans, LA, 68-73, 1991.

Molinari, J., and M. Dudek, Parameterization of convective precipitation in mesoscale numerical models: A critical review, *Mon. Weather Rev.*, *120*, 326-344, 1992.

Nickerson, E. C., E. Richard, R. Rosset, and D. R. Smith, The numerical simulation of clouds, rain, and airflow over the Vosges and Black Forest mountains: A meso-beta model with parameterized microphysics, *Mon. Weather Rev.*, *114*, 398-414, 1986.

Paoli, G. M., *Climate Change, Uncertainty and Decision-Making*, 164 pp., Institute for Risk Research, University of Waterloo, Waterloo, Ontario, Canada, 1994.

Sass, B. H., and J. H. Christensen, A simple framework for testing the quality of atmospheric limited-area models, *Mon. Weather Rev.*, *123*, 444-459, 1995.

Trenberth, K. E., J. C. Berry, and L. E. Buja, Vertical interpolation and truncation of model-coordinate data, *NCAR Tech. Note.*, *TN-396+STR*, 54 pp., National Center for Atmospheric Research, Boulder, Colo., USA, 1993.

Waldron K. M., J. Paegle, and J. D. Horel, Sensitivity of a spectrally filtered and nudged limited-area model to outer model options, *Mon. Weather Rev.*, *124*, 529-547, 1996.

Zhang, D., E. Hsie, and M. W. Moncrieff, A comparison of explicit and implicit predictions of convective and stratiform precipitating weather systems with a meso-Beta-scale numerical model, *Q. J. R. Meteorol. Soc.*, 114, 479, 31-60, 1988.